Exploring Kawayan Dialect Lexicon (Wordlist) in Southern Negros, Philippines

John Gerald Arbias Pilar[1]* (iD)

[1] General Education Department, Carlos Hilado Memorial State University, Bacolod City, Philippines

A R T I C L E   I N F O

A B S T R A K

Luasnya penelitian leksikal mengakibatkan pengembangan daftar kata apa harus bergantung pada pertimbangan kode-glosses dan kata-kata alternatifnya untuk memahami maknanya dan status bentuk-bentuk tersebut. Penelitian ini mencoba mendeskripsikan leksikon dialek Kawayan dari 278 daftar kata yang lebih mudah ditelusuri ke keluarga Bisayan dan tampaknya adalah Bisayan Barat dengan memanfaatkan dan memanipulasi Pivot Excel Table untuk menelusuri cap waktu untuk mencabut kata-kata dari informan (penutur asli) . Penelitian diserahkan untuk pemeriksaan silang antar antar pembuat kode untuk dilihat dan diperiksa validitas dan reliabilitasnya sebagai protokol untuk memverifikasi keakuratan kumpulan data dari penutur asli. Daftar kata ini penting dalam memberikan informasi tentang hubungan genetik dialek Kawayan dengan kelompok bahasa Philippines. Selain itu, menggali lebih dalam tentang bagaimana penelitian ini disajikan dan disebarkan kepada pembaca dan cendekiawan yang dituju merupakan hal yang produktif dan membuka mata. Karena ini adalah penyelidikan dan dokumentasi yang dirintis terhadap dialek Kawayan, meskipun mungkin menjengkelkan, setidaknya ini dimulai sebagai salah satu ciri khas dari dialek lain yang belum termasuk dalam daftar dialek Filipina. Peneliti saat ini mendesak kita untuk mengeksplorasi dan menyelidiki dialek ini sebelum hilang tanpa jejak.

A B S T R A C T

The extent of lexical research has resulted in the development of word lists, which have to rely on consideration of code-glosses and their alternative words to understand their meaning and the status of these forms. This research tries to describe the Kawayan dialect lexicon from a list of 278 words that are more easily traced to the Bisayan family and appear to be Western Bisayan by utilizing and manipulating an Excel Pivot Table to trace timestamps to extract words from informants (native speakers). The research was submitted for cross-checking between coders to be reviewed and checked for validity and reliability as a protocol for verifying the accuracy of data collection from native speakers. This word list is essential in providing information about the genetic relationship of the Kawayan dialect to the Filipino language group. Additionally, digging deeper into how this research is presented and disseminated to its intended readers and scholars is productive and eye-opening. Since this is a pioneering investigation and documentation of the Kawayan dialect, although it may be annoying, at least it started as one of the distinctive features of other dialects yet to be included in the list of Filipino dialects. Researchers today urge us to explore and investigate this dialect before it disappears.

1. INTRODUCTION

Recently, there have been attempts to study the local dialects in the Philippines, making another milestone in exploring the vanishing dialects. One is the Kawayan dialect. This is one of the local dialects in Negros Island. Thus, attempting to explore its lexicon or wordlist is the focus of this investigation. Kawayan dialect is one of the dialects in the archipelago, which still needs to be completed (Endardi et al., 2023; Tondo, 2020). It is presumed that Kawayan dialect uses by the locals in daily conversations in the Southern part of Negros Occidental but undocumented and under study. This Kawayan dialect may exist in the municipality of Cauayan and other neighboring areas, namely; Cauayan, Hinoba-an, Ilog, Candoni, Kabankalan City, and Sipalay City in the 6th district of Negros Occidental and the highland parts between Negros Oriental and Negros Occidental (Hirota & Tsuji, 2021; San Jose & A. Pilar, 2023).

*Corresponding author
E-mail addresses: john.pilar@chmsu.edu.ph (John Gerald Arbias Pilar)

On the other hand, this study clarifies the term 'dialect,' which alternately means 'language,' and inventory also means a lexicostatistical database of sounds. A thorough phonological description will also describe the variations of the sounds and when and where to make them adequately understood when speaking or hearing with understanding (Maharani et al., 2023; Menn & Stoel-Gammon, 2017). The lexicon or the wordlist develops in this study to have initial exploration and documentation to aid in developing Kawayan vocabulary. This Kawayan lexicon may assist with the process of day-by-day conversation of the native speakers that often occurs throughout different situations. A wordlist should be at the core of effective vocabulary creation and adequate vocabulary resources in high-frequency language learning (Burkett, 2015; Dang et al., 2022). Specifically, the Philippines has more indigenous languages than non-indigenous ones. There are still many linguistic places in the country that still need to be explored or adequately examined. There is no objective and scientific way of identifying when to use the words 'language' and 'dialect' (Melinger, 2018; Sherif et al., 2023). Distinguishing languages from dialects depends on at least three factors: mutual intelligibility, the speakers' culture or opinion, and political status (De la Torre & Gonong, 2020; Mahmud et al., 2023). The number of linguists who have exceeded those 200 or so essential vocabulary items has increased in recent years, thereby increasing the repertoire of forms that comparativists have to work with. With so many researchers now in the field and with a growing concern over the linguistic situation in Borneo and Celebes (many of the languages of which are close in type to languages of the Philippines) (Blust, 2019; Enfield & Comrie, 2015; Sohayle et al., 2022), the time is ripe for a serious attempt at drawing up a definitive, enlarged, culturally-oriented list in order to ensure a larger body of semantically matched data for continued comparative work on the one hand, and in order to begin the reconstruction of Philippine culture history on the other.

Many words are not in common use across a complete language area, and indeed, many terms of a language are known only to speakers of a limited area, it is also often difficult to recognize the phonetic qualities of adaptations. from a particular language (Kleinschmidt & Jaeger, 2015; Lewandowski & Jilka, 2019). The realization and detailed observation of common words often have distinct pronunciations in local speech and meanings that differ regionally, and language development is the study of linguistic features that differ systematically among different groups of speakers or the same speaker in other contexts (De la Torre & Gonong, 2020; Symaco, 2017). In this premise, the Kawayan lexicon is the focus of this study on what words are in the locality and what meanings are assigned to them. This study will also be an impetus to a lexicostatistical database of the said dialect on the dialect lexicon of the Kawayan dialect. The collection of words and information is on the vocabulary used in one and at a specific time and their assigned and accepted glosses within the speech community.

Specifically, this study focuses on the lexicon of the Kawayan dialect. They are categorizing lexical items from the utterances of the native speakers in the locality. This study will also describe the current dialect contact situation within the speaking area of the Southern Negros, and illuminate the historical, social, and cultural factors that have existed or such a variety arise in analyzing the data. This study will further be the basis of a lexicostatistical database of the lexicon of the Kawayan dialect to be used in revitalization, preservation, teaching MTB-MLE using the word list, and future researchers in the field. So, the desire to scientifically give students, teachers, researchers, and the Local Government Unit (LGU) significant resource material and literature in studying the Kawayan dialect lexicon provides reference, revitalization, preservation, and realization as a reflection of the true identity of southern people in Negros Occidental through examining the dialect lexicon of Kawayan dialect.

2. METHOD

This study employed a qualitative text analysis method. This approach is a resource-intensive process, more diverse and complex from the data that may contain transcripts of face-to-face interviews and other interactions between researcher and informants (native speakers). Moreover, the researcher limits to smaller sample sizes when analyzing text data, assigning code labels, and iteratively developing findings (Guetterman et al., 2018; Smith, 2017). This study assumes that reality emerges through language so that researcher intrusion would alter the creation of truth in an unnatural direction. Hence, detached observation is required. Thus, this study focused on Kawayan dialect lexicon analysis through linguistic modeling, semantics meaning, and code glosses. This study is qualitative by nature. Transcriptions from the interviews and audio recordings were the basis for collecting the raw data. The researcher must carefully consider how their position and perceptions might influence the data collected and explicitly address the concerns in collecting the raw data through interviews. Thus, the researcher asked an expert in the field to assist the former. The interview protocol flow enables the claims of the research and determines its generalizability or transferability (King & Mackey, 2016; Larsen-Freeman, 2018). The researcher's interview protocol presents his experiences in the field.

The informants were the local/native speakers of the following places namely: Candoni, Cauayan, Hinoba-an, Ilog, Kabankalan City, and Sipalay City (CCHIKS) of the southern parts of Negros Occidental, Philippines which purposive sampling technique was employed in this study. There were conditions, limitations, and willingness of the informants to participate in the interview because of the pandemic concerns. The researcher took only the three areas of the southern parts of Negros Occidenal such as Cauayan, Hinoba-an, and Sipalay City. There were 9 informants participated in the study. 3 (females) in Cauayan, 1 (male) in Sipalay City, and 2 (males), 3 (females) in Hinoba-an who informed and voluntarily participated in the interview. The informants narrated their personal stories and folktales in the locality since the primary purpose of the study is exploring the undocumented lexicon of the dialect in the locality. Thus, no thematic insights and themes were formed and analyzes but the frequency of the lexicon from the interviews' responses to create data set.

The researcher adopted selection criteria (De la Torre & Gonong, 2020; Malmasi & Dras, 2017) and made modifications to filter the target informants. In this procedure, the researcher strictly followed and observed the selection criteria with the consultation with the some experts in the field in regards with the researcher's informant: (1) A native/local speakers and had lived more than 50 years in that area; (2) He/she should be more than 70 years old and preferably 100 years old; (3) He/she is a well-versed in their native/local tongue/dialect; (4) He/she could tell folktales aside from their personal life stories. This study utilized and manipulated an Excel Pivot table to trace the time stamp for retracting the words from the informants (native speakers). It was further submitted for cross-examinations among inter-coders which they were looked and checked the validity and reliability as protocol of verifying the accuracy of the data set.

Thus, this study selected qualified inter-coders to establish the validity through inter-coding. A thorough discussion among inter-coders met to develop reliable results by resolving the boundaries of the analysis process. Each objective of this study scrutinizes the inter-coders—redefined agreement among inter-coders on the concepts needed to attain the agreement of the majority (Corcuera & Bernardo, 2024; Madrunio, 2022). This study adopted the protocol in fieldwork in the locality for almost 11 months. There were various revisions on the protocol done in this study specifically in gathering the data set for final analysis. The data gathered link to a given speech variations of at least two for example, historical narratives or folklores of their selection and autobiographical or personal stories (Tamano et al., 2021; Zorc, 1974). The adjustments of the researcher in collecting the data and fieldwork were reflected and documented: (1) show the purpose and aim of the study to the Local Government Unit (LGU); (2) collaborate with the local assistants and institutions or organizations in the area; (3) recognizing the shortlisted informants usually aged 60 and above but preferably 100 years old; (4) road mapping of the areas of the shortlisted informants; (5) asking consensus and willingness of the informants as well as their immediate family member(s) by explaining the purpose of research; (6) there are two settings of collecting data such as asking they are autobiographical and to tell folklores which handed down from their ancestors; (7) there will be revisiting of the identified informants for validation and confirmation of their stories like autobiographical and folklores using electronic gadgets in capturing the voice narratives of the informants.

It was distinguishing the approach used in similar studies, the phono-lexicostatistics method in the data analysis comprised of phonetic and lexical identification, which includes explicitly defined criteria adapted from Blair (De la Torre & Gonong, 2020; Malmasi & Dras, 2017). The present researcher modified the approach used by the previous authors, as mentioned earlier. These modifications are the 278-word list, which is composed of (1) words that are at least present from 2 out of 9 informants, (2) the inter-coders and intra-coder had selected and further validated by taking consensus among them the identification of the content words whether they are nouns, verbs, adjectives and adverbs. The inter-coders used the syntactic rules and core grammatical categories to select the word list for the reliability of the data for interpretation. In examining the Kawayan dialect lexicon, the units of analysis utilized the word level for linguistic modeling and code-glosses, while phrasal and clausal level analysis was for the semantic meaning.

3. RESULT AND DISCUSSION

Result

The study attempts to describe the Kawayan dialect lexicon of a 278 wordlist. A word is composed of form (a cluster of phonological features) and meaning (a cluster of semantic units). The gamut of lexical research deals with each aspect. In fieldwork, glosses (usually translations in a contact language) are given to elicit forms that accurately translate those meanings. The development of any wordlist must consider the glosses that should be used, the forms expected to be elicited by those glosses, and the status of those forms.

Table 1 is 278 wordlist of Kawayan dialect lexicon forms that are more readily traced to the Bisayan family and appear to be West Bisayan. This wordlist is significant in giving information about the genetic

relationship of the Kawayan dialect to a Philippine language group. Code glosses are used to elaborate on what has just been said and gain the appropriate meanings of elements of the narratives. Sometimes, it defines or explains a word, phrase, or idiom, signaling the rewording of knowledge regarding the ideation (Alqahtani & Abdelhalim, 2020; Nausa, 2019; Safari, 2018). While the 'alternate words' are semantically related to the word best completing the sentence or closely related language or word (Artetxe et al., 2018; Kutas et al., 2019). In this study, code-glosses and alternate words are included in the Kawayan words wordlist for future reference, and it may help researchers further explore and investigate.

Moreover, digging deeper into how this study was presented and transmitted to the intended readers and scholars is prolific and eye-opening. Since this is a pioneered investigation of the Kawayan dialect, although exasperating as it may be, it at least started as one of the hallmarks of dwelling another dialect that still needs to be included in the Philippine list of dialects. The present researcher is now urging to explore more on this present study.

Table1. Kawayan 278 Wordlist

No.	Kawayan Words	Code-Glosses	Alternate
1	alagyan	passage	way
2	alihid	side	corner
3	amomonggo	legend creature	Ape-like Humanoid
4	anggid-anggid	same	look-alike
5	angkunon	take	grab
6	anibong	kind of leaf	
7	aroy-aroy	weak	sick
8	asawahay	married	husband and wife
9	asawahon	wife	spouse
10	babaylan	faith healer	Shaman
11	bagakay	bamboo*	
12	bahag-bahag	half-naked	dance or ritual
13	balati-an	illness	sickness
14	balatik	baton	bamboo stick
15	bantayan	monitor	watch
16	binangon	machete	sword
17	binulan	monthly	once a month
18	binulod	anthill	legend- house of mystical dwarfs
19	bukid-bukid	mountainous	forest
20	bulawan	golden	legend- golden cow
21	butang-butangan	allegations	deceive
22	butangan	place	put
23	Camindangan	name of a place	
24	Candalamoan	name of a place	
25	Cansibit	name of a place	
26	Cansinudlay	name of a place	
27	daga-daga	minor offering	bait
28	dagyaw-dagyaw	common labor	collective work
29	dalagko	big	huge
30	dalawat	cuy rice	
31	diwata	fairy	nature spirit
32	diwatahan	ritual offering	ritual for nature spirits
33	duludalagko	slightly bigger	medium
34	dumduman	remembered	recalled
35	durgoson		
36	duslakon	stab	
37	engkanto	enchanted	nature spirit
38	gabahag	wearing g-string	
39	gadaig	burning dried leaves	setting fire
40	gahagunos	galloping	swift
41	gahali	repeatedly done	chronic
42	gahambal	speaking	talking
43	gahampang	Playing	
44	gakaingin	slash and burn farming	

No.	Kawayan Words	Code-Glosses	Alternate
45	gakatabo	happening	occurring
46	gakatum-ok	pressed at	put-down
47	galibot-libot	Circling	roaming around
48	ganiwang	becoming think	losing weight
49	gapakigbagay	mingling with	being one of the boys
50	gapamintas	being cruel	violent
51	gapamulong	Healing others	healer
52	gapanagat	going to sea	catching fish at sea
53	gapanago	Hiding	covering at/under
54	gapanaka-saka	gate-crashing	burglary
55	gapananum	planting	farming
56	gapangaon	eating	devouring
57	gapangguba	act of destroying	breaking the something
58	gapanghinas	collecting shells	
59	gin-hingadlan	named	christened
60	ginakilala	being honored	known as
61	ginasiling	said to be	rumored as
62	ginikanan	parents	old folks
63	haladan	altar	offering table
64	haom-haom	assumed	conjecture
65	hapa-hapa	laying on stomach	
66	hawot-hawot	lacking	insufficient
67	hidlawon	being missed	
68	hinablos	nephew or niece	
69	hinambak	way of speaking	
70	hinambalan	language	agreement
71	hinangpanay	common understanding	unanimous
72	hinangyawan		
73	hingadlan	to name	will be called
74	Hinobaan	name of a place	
75	hubin-hubin	slightly younger	very young
76	hunasan	beach	shore
77	ibalik	to return	bring back
78	idalawat	will buy rice	
79	ilaya	highland	upland
80	inambak	way of speaking	
81	ipakompras	to sell at a whole sale price	
82	iunong	stay with	standby
83	Kabankalan	name of a place	
84	kabati	heard from	can hear
85	kabudlay	difficult	effort
86	kabuhi	life	
87	kabukiran	mountains	upland
88	kadalagko	big in size	
89	kadamo	number	population
90	kadukasan	area	place
91	kagapon	yesterday	the day before
92	kagulangan	forest	
93	kahibalo	known	knows
94	kahigkahigon	brush away	
95	kahiyabaon		
96	kaingin	slash and burn	
97	kakuyangan	jungle	woods
98	kalabanan	mostly	oftentimes
99	kalasangan	forest area	
100	kalatian	new moon	
101	kalilibog	mindboggling	
102	kaliwatan	race	

No.	Kawayan Words	Code-Glosses	Alternate
103	kalolalolahan	great great grand mother	
104	kalolalolahan	great great grand mother	
105	kalolohan	male grand parents	
106	kalomboyan	*Lomboy* tree orchard	Native Plum
107	kamalig	Shanty	Nipa Hut
108	kamangyan	kind of vine	Panyawan vine
109	kamatuoran	truth	fact
110	kanalang		
111	kanami	nice	good
112	kapigaduhon	poverty	economically challenged
113	kapuraot	difficulty	scarcity
114	kasami	frequently done	oftentimes
115	kasiot	uncleared area	bushes
116	kasubo	lonesome	sad
117	kataas	high	height
118	katanda	remember	sharp eye
119	katapusan	ending	last
120	katigulangan	elders	parents
121	katingalahan	wonderment	amazement
122	kaugalingon	self	own
123	kawatan	thief	burglar
124	Kilat-an	being strike with thunderstorm	being cursed
125	kinaiya	natural characteristics	identity
126	kinalain	difference	unique
127	kinalasag	ritual	dance for the festival
128	kinamatarong	rights	dignity
129	kinamatarung	rights	dignity
130	kinaugali	tradition	practices
131	kinaugalingon	of one's self	culture
132	lain-lain	different	seperate
133	laka-laka	scarce	spaced at
134	lamagkaw		
135	laminpaw		
136	lansangan	pinned with nail	to put nail
137	lapit-lapit	closely	look similar
138	lasang-lasang	minor forest	remote area
139	likit-likit	rolled	slightly broken
140	linagupok	unusual sound	
141	lingaw-lingaw	entertainment	
142	lubi-lubi	like a coconut	
143	lutsanan	passage	
144	luyaganay	attracted to each other	in love
145	maayad-ayad	slightly good	
146	mabugyawan		
147	mabulos	to substitute	will replace
148	madanaw	water will spill	
149	maglabay-labay	passed-by	
150	magluyak		
151	magutod	brothers/sisters	
152	mahagbas	to cut grasses	cut leaves by slashing
153	makabati	can hear	
154	makadawat	can receive	
155	makakadlaw	can laugh	
156	makakita	can see	
157	makipagpatay	to kill	fight to death
158	manadi	legend of the demi-god	legendary of a supernatural hero
159	managat	went to sea	go fishing
160	mangakig	will get angry	

No.	Kawayan Words	Code-Glosses	Alternate
161	mangaman	to prepare	to be ready
162	mangasawa	to get married	to get a wife
163	mangawat	to steal	to rob
164	manggaranon	tich	affluent
165	mansig-mansig	on his/her own	on one's own terms
166	mapalapit	will get close	
167	mapisik	agile	quick in action
168	Maricalum	name of a place	
169	mariit	enchanted place	controlled by bad spirits
170	masakiton	frail	sickly
171	masaot-saot	will dance	ritual dance
172	masarangan	can overcome	Can bear
173	masumpa	will swear	will take oath
174	matapos-tapos	close to finish	not yet done
175	matuod	truth	real
176	matuod-tuod	very true	so real
177	maunong	will standby	will stay
178	minahan	mine	
179	naabtan	reached	catch up
180	naabutan	on-time	
181	Nabulaw	name of place	
182	nag-uyat	hold on to	gripped at
183	nagalakot	mixed up	
184	nagalanog	heard in wider area	extremely loud
185	nagapabaybay	going to the beach	
186	nagapadayun	continuing	on-going
187	nagkalamatay	died	extinguished
188	nagkilit-anay	met each other	converged
189	nagkilitaay	crossed path	
190	nagpahibalo	announced	informed
191	nagtabangay	helped each other	back up
192	nahadlok	afraid	scared
193	nakadalaw	visit	
194	nakakilala	knows	remembered
195	nakapamana	got married to a husband	
196	nakapangasawa	got married to a wife	
197	nakapuli	got home	
198	nalagpatan	unexpectedly	coincidence
199	namana	got married	
200	namangkot	asked a question	
201	namatyan	bereaved	
202	nanga-matay	died	
203	nanigulang	get older	aged
204	napamana	married to a man	
205	napiskan	ricocheted off	
206	nasakupan	territory	kingdom
207	natabuan	coincided	
208	natawgan	called by	reached out
209	natubuan	grown with	being matured
210	Pachucho	name of a place	
211	pagangabuhi	living by	surviving
212	pagka-sinabawan		
213	pagkaingon	said	response
214	pahuway	rest	relax
215	pamalatik	laying trap	
216	pamana	get marry to a men	
217	pamanaha	marry someone	
218	pamatyagan	feeling	impression

No.	Kawayan Words	Code-Glosses	Alternate
219	panabog	send away	
220	panagat	going to sea	livelihood being fisherman
221	pananom	crops	planting
222	pangabuhian	means of living	job
223	pangawat	stealing	robbing
224	pangayaw	stranger	traveler
225	pangayawat	thrifty	economical
226	panghambal	manner of speaking	utterances
227	panghusay	combing of hair	mediator
228	panguma	farming	
229	panimalay	household	home
230	panimuot	consciousness	understanding
231	panulay	tempter	devil
232	panulok	vision	sight
233	panutod	burning of dry leaves, etc.	
234	papag-papag	makeshift bed	ritual to the bad spirit
235	pasaka	scending	reclining
236	pasibangdan	accused of	
237	patadyong	kind of cloth use as a skirt	
238	patok-patok	hitting bamboo to make noise	making sound to a ritua
239	patubas	crop	yield
240	pinanghambalan	informed	advised
241	puti-pution	slightly white	fair complexion
242	sagawid	blocking	hindrance
243	salalakot	mixture of various parts	mixed up
244	saot-saot	dancing	rejoicing
245	sapa-sapa	creek	small river
246	singganan	told	informed
247	sinulog	a kind of dance	ritual dance
248	sugid-sugid	rumored	news
249	sugilanon	story	language
250	tablitan		
251	tag-lugar	enchanted spirit	
252	tago-tago	hidden	hiding playfully
253	tagsubong	of the same kind	at this point
254	tahod-tahod	minor spur	respect
255	talagsa	once in a while	not often
256	talamnan	field	farmland
257	tamaran	lazy	
258	tamawo	enchanted spirit	
259	tanan-tanan	all	everything
260	tanom-tanom	leisure planting	
261	tig-ulolan	rainy	
262	tiligangan	container used to cook rice	pot
263	tinaga	word	
264	tinuig	years	
265	tipon-tipon	getting together	saving
266	tiyog-tiyogan	a space to move-around	to spin
267	tribu-tribu	tribe	
268	tuburan	spring	
269	tulisan	marauders	bandits
270	tulutaklad	slightly ascending	
271	tumandok	native	locals
272	ugangan	parent-in-law	parents of one's spouse
273	ulingan	place to process charcoal	
274	umagad	son/daughter-in-law	
275	updanay	accompany	colleague
276	uru-ibabaw	slightly on top	

No.	Kawayan Words	Code-Glosses	Alternate
277	waswasan	strike in full force	beaten
278	yagutaon	socked at	bully

In the previous studies on the most distinguished were the lexicostatistical lists/inventories of Swadesh (100 and 200 glosses) and Gudschinsky (215 glosses). These have been labelled on the notion of containing and gaining collective language meanings that are non-cultural. Researchers have utilized these lists/inventories in every reliable creation for example, it was alphabetically organized n the linguist's native or local tongue and a contact language or even the business language. It also arranged in semantically by clusters for instance body parts, pronouns, natural experiences, etc. and it further arranged accordingly to the proposed etyma for each gloss. Structural grouping, i.e., the arrangement of items by word classes (noun, verb, adjective), among some lists for use in the New Guinea area, but no such arrangement has yet been noted among lists for use in the Philippines. In this case, the 278 wordlist of the Kawayan dialect was meticulously selected by intercoders according to the selection criteria presented.

Based on Table 1, due to the variation of vocabulary demands, classifying depends on the narrative of the native speakers/informants. For example, the informants narrating his/her biography will be marked in another category of vocabulary lists than the folkloric narratives of the informants because there are word families that cannot reach the same threshold being analyzed. Another noteworthy detail was the proportion of affixes in the word, which may change accordingly, demonstrating the relative importance of recognizing and understanding by recognizing its root word to proceed with code-glosses and alternate words.

It is further displayed two hypothesized scenarios: one assumed that all Kawayan words have affixes and how about reduplication of root word, for example, *dagyaw-dagyaw* (help one another), *saot-saot* (dancing), *tipon-tipon* (getting together). This reduplication of root words may be categorized in another vocabulary list of the Kawayan dialect. The second scenario is lexical demand for more tokens in combination, and the lexical profile of the corpus deserves a deeper investigation into its variation in the data for such analysis. The present data limit for these coverage thresholds is still worth nothing because the Kawayan dialect depends on affixes in determining the meaning of each word and its alternate word. Reduplication of the root word is obligatory to utilize hyphenation, which is also observable in this dialect (San Jose & A. Pilar, 2023; Symaco, 2017).

Nevertheless, it is not a one-size-fits-all observation of this investigation since there are words in the wordlist with no code glosses or alternate words. However, they are found in the utterances of the informants, which were qualified among intercoders based on the selection criteria. These are included in the wordlist for future researchers to explore these words and their meanings. It is also subject-specific vocabulary, and term recognition suggests that quantitative and qualitative selection criteria are necessary to identify these words reliably. In this view, transcribed words representing minor linguistic family like the Kawayan dialect, which is still an undocumented dialect of the Philippines, has uncommon words and do not frequently occur even in the nearest Bisayan languages like Hiligaynon, Waray-Waray, Cebuano, etc., across Bisayan languages and families. This new evidence suggests that the role of ease of documenting and exploring the Kawayan lexicon may be more influential than generally assumed. However, other argument factors are also likely at work, and these structures require further exploration with more robust intralinguistic lexicon corpora (Everett, 2018; Jablonkai, 2020).

On the other hand, it is assumed that a reasonable and acceptable number of wordlists than previous studies would result in acceptable analyses. May be interesting since understanding the locals could be a great way to establish a sense of connection and belonging to a country's local communities and networks (Gomes, 2015; Juang et al., 2018). It is conservative and wealthy as teaching materials for the mother-tongue teachers. It could be an excellent source for teaching the local language as it provides up-to-date and interesting information while offering rich linguistic resources, such as folklore. Generally, input resources selected for language learning should be lexically less challenging than in the real world. Hence, teaching language through wordlist could be proficient in learning the language, which is the number one concern for lesson and material design for learning and could also be put into consideration when choosing lexical words to teach to the young generation so that it will remain and inculcate to their vocabulary as Kawayan dialect is concerned. Focus on learning the vocabulary that is specified and specialized and direct vocabulary learning to more specialized areas when learners have mastered general usefulness (Ha, 2022; Pathan et al., 2018). The following sample extracts are the semantic meaning and lexico-grammatical features of the Kawayan lexicon to show its structures and forms:

Extract 1

> …Anu ni kay **mangasawa** na kag ang among anak kwan kabilog pulo*…. (This is [about] to marry and then our son hmmm… they are ten (10) siblings….)*

The lexical word *mangasawa* is defined as marrying a man (husband); it is a woman wanted, and she belongs to 10 siblings. Looking at the semantic meaning of this extract 1 and its English nearest translation, it may be determined that *mangasawa* is from a root word- *asawa,* which means in Hiligaynon-wife. This is also true in the Kawayan dialect, where the term *asawa* means wife. Unlike the term *mag-AsaWA*, it means couple. Still, the root word is *asawa.* In Tagalog language, *asawa* means for both genders, either husband or wife. Generally speaking, *asawa* means a couple.

Extract 2

> *…*Tiya ko **nakapangasawa** sang siling nila bala nga Bukidnon. Native diri sa Maricalum*….([My] aunt married [they] said here that's it a (Bukidnon- a native tribe). Native here in Maricalum….)*

In extract 2 lexical word, *nakapangasawa* is defined as married to a native man. It is obvious to say that married to a man because it signals the word *tiya (aunt)*. Pragmatically, *tiya* married a native man in Maricalum (a place in Sipalay City, Negros Occidental, Philippines).

Extract 3

> *…Ang amay ko nag* **_asawahay_** *sila diri…. ([My] father, [they] married here….)*

Extract 3, *asawahay* lexical word manifested married couple. It observed in extract 1 that the definition of *asawa* is a couple in the general speaking scenario; likewise, in the Tagalog language *asawa* is intended for both genders. Indeed, this lexical term has universal meaning in most Philippine languages and indicates the interrelationships among languages in the Bisayan languages and the entire Philippine archipelago. The Philippines possesses a great wealth of indigenous languages, which are related. Moreover, some words have a degree of closeness to these languages, which are identical in vocabulary items.

A glance at the sample extracts 1, 2 and 3 will show how lexico-grammatical features are demonstrated in table 2.

Table 2. Sample Lexico-Grammatical Features

Lexical Item	Root Word	Syllabication	Affixation
Mangasawa	asawa	manGAsawa	*mang-*
Nakapangasawa	asawa	nakapanGAsawa	*naka- + pang-*
Asawahay	asawa	asaWAhay	*-hay*

*capitalization of the syllable means stressed otherwise unstressed for not in capital

Table 2 observes the structure of the syllabication that stressed syllables are usually found in the middle of the word, and it sounds like a 'glottal stop.' This is a characteristic of the Kawayan dialect of nasality of the sound system regardless of affixation attached to the root word. Moreover, the affixation found in Table 2 exhibits verb form, for example, *mang-, naka-, pang-* as prefixes and *-hay* for the suffix. In this view, the Kawayan lexicon is governed by affixation to create a new form of the word and its function in the utterances. However, the present researcher cannot determine the degree of frequency of the lexical features, specifically the affixation of the Kawayan lexicon, because of the limited tokens to justify the argument based on the data set. This is now the challenge for future researchers and experts to continue exploring the Kawayan wordlist given the difficulty of creating and selecting a wordlist of the undocumented dialect, for instance, the Kawayan dialect. Given the assumption that frequency and coverage are essential criteria for selecting a wordlist or collection of texts, it is a valuable source of empirical information that can be used to examine the language in depth. However, how it should be developed needs to be clarified, and the criteria used for word selection.

Discussion

Not all words in the 278 wordlist have code glosses, as shown in the data, but fewer frequency counts appeared. A possible explanation might be that there is a limitation in categorizing the narratives from the informants, which are supposed to focus on folkloric or personal life stories. In other words, an interpretation for this might command a description of what is reflected in the wordlist, as this study's

objective is to explore the Kawayan dialect lexicon (Safari, 2018; Saimon et al., 2021). Taking into account that this study has presented the 278-wordlist which examines the Kawayan lexicon in a more panoramic view to obtaining at least a tentative conclusion that would be interesting to future experts in the field to investigate another angle of this Kawayan lexicon since exploring this study is clear from the start of the discussion, it is experienced the present researcher to mention the gap by transmitting to the future researchers to dig deeper into the complex linguistics structure of the Kawayan dialect.

This wordlist has a quality impact on the local communities in learning materials in teaching the mother tongue. This information presentation is crucial to its legibility and potential impact on the list of dialects in the Philippines to be included. In this way, the agency handling this list of dialects can map and document to preserve and revitalize, considering the fast evolution of digitalization and language change in the new millennium. Thus, the vanishing of this dialect is inevitable and a permanent reality nowadays because digital platforms could influence learning a new foreign language rather than a local dialect. Despite the limitation, the findings have several significant implications, one of which is the value of attempting to explore the undocumented Kawayan dialect by incorporating technology in collecting data from informants and application of pivot table in Microsoft Excel application as a text analyzer before submitting to the experts (intercoders). Furthermore, it provides a simple tool to assist the present researcher in improving the accuracy and validity of the data set towards the wordlist. In this manner, future researchers may guide how to employ technology in this research protocol to increase data reliability for analysis.

Indeed, it is an indispensable part of language learning, and researchers agree that there needs to be more research carried out on local dialect exploration and discovery. In recent years, applied linguists have strongly advocated vocabulary teaching, in this case, the Kawayan wordlist, because of the vital role it plays in language learning by helping in creating wordlists from different contexts, for instance, personal stories of the native speakers (informants) or telling folklores within the community. Understanding how words are organized and how word meanings are connected in a native speaker's lexical repertoire becomes increasingly essential to teaching the learner's lexical repertoire ability to acquire new vocabulary, such as the Kawayan lexicon. There are words that intercoders cannot decode and their meanings and applications because of the limited semantics and narratives from the data set. Thus, the reflection of the quality of the wordlist depends on the semantic meaning and lexico-grammatical feature of the dialect, which leads to the tentative conclusions of this exploration and investigation of this Kawayan lexicon.

In learning a new dialect, for example, the Kawayan wordlist utilizes the decontextualized activities of the vocabulary elements, which are removed from the context in which they first appear and are presented in situations free from any communicative values. However, it is recommended to utilize fully contextualized activities that are supposed to play an important role in vocabulary teaching. They provide students opportunities to practice authentic communication in real-life situations that may encounter multiple varieties of words in the target dialect, like Kawayan. The Kawayan wordlist presented in this present study, along with code-glosses and its alternate word to one side of each word, is the technique which many experts believe that working with a wordlist is one of the most effective ways of acquiring new language vocabulary from the complex nature of words, learners need to encounter words in several meaningful contexts. It is now a challenge for future researchers to grab the opportunities to dig and explore further by making connections between what they already know in some other Bisayan languages like Hiligaynon, Waray-Waray, etc. Although a great deal of work has been done on the Kawayan wordlist, and somehow it may seem a little work has also been done on exploring the Kawayan lexicon, it is a bridge to fill the gap on undocumented dialect in the Philippines and addressing the gap of the scarcity of the mother-tongue instructional program.

4. CONCLUSION

The native speakers have distinct accents and vocabulary in assigning its meaning or, in other contexts, with different purposes and usability in the Kawayan dialect, so much more in the lexical aspect. Thus, a rich vocabulary in the dialect reflects how distinct and diverse some words exist in the dialects. The manifestation of the Kawayan dialect and its lexicon exhibit a simple but rigorous hybrid analysis approach from word concordance using an Excel pivot table and employing inter-coders (experts in the field) to strengthen the accuracy and consistency of the data and results. This study serves as the foundation for exploring the Kawayan dialect lexicon in Southern Negros, Philippines, by capturing the lexical elements, but still, further investigation and exploration of the remaining morphosemantic roots are the opportunities to discover. It is vital that any teaching and learning of vocabulary from a wordlist needs to be done in a principled way, for example, the Kawayan lexicon, which contains a discussion on acquiring new words towards vocabulary program through mother-tongue-based education. Therefore, there should be equal opportunities for learning vocabulary, and teaching language activities should actively engage learners in

working with target vocabulary that matches learners' readiness and preparedness, like learning the Kawayan lexicon using the developed wordlist of this present study. However, learning new vocabulary of the Kawayan lexicon might never find the words in context in the materials they are reading and might never practice the words meaningfully. For example, learners might focus only on the spelling and meaning of words rather than on using the words themselves in speaking and writing since the Kawayan dialect is new to them.

5. ACKNOWLEDGE

6. REFERENCES

Alqahtani, S. N., & Abdelhalim, S. M. (2020). Gender-based Study of Interactive Metadiscourse Markers in EFL Academic Writing. *Theory and Practice in Language Studies*, *10*(10), 1315. https://doi.org/10.17507/tpls.1010.20.

Artetxe, M., Labaka, G., & Agirre, E. (2018). A robust self-learning method for fully unsupervised cross-lingual mappings of word embeddings. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 789–798. https://doi.org/10.18653/v1/P18-1073.

Blust, R. (2019). The Austronesian Homeland and Dispersal. *Annual Review of Linguistics*, *5*(1), 417–434. https://doi.org/10.1146/annurev-linguistics-011718-012440.

Burkett, T. (2015). An Investigation into the Use of Frequency Vocabulary Lists in University Intensive English Programs. *International Journal of Bilingual & Multilingual Teachers of English*, *03*(02), 71–83. https://doi.org/10.12785/ijbmte/030202.

Corcuera, J. K., & Bernardo, A. S. (2024). On-Air, Off-Balance: Detecting and Examining Media Biases in Filipino Sportscasting. *Linguistics Initiative*, *4*(1), 88–120. https://doi.org/10.53696/27753719.41142.

Dang, T. N. Y., Webb, S., & Coxhead, A. (2022). Evaluating lists of high-frequency words: Teachers' and learners' perspectives. *Language Teaching Research*, *26*(4), 617–641. https://doi.org/10.1177/1362168820911189.

De la Torre, R. G., & Gonong, G. O. (2020). A phono-lexicostatistical analysis of Bikol-Sorsogon varieties. *The Normal Lights*, *14*(2). https://doi.org/10.56278/tnl.v14i2.1655.

Endardi, J., Gunadi Kasnowihardjo, M., Sukesti, R., Kurniawati, W., Sa'diyah, U., Triyoko, H., & Atmawati, D. (2023). Relics And Innovation In The Bawean, Sampang, And Rembang Dialects (A Collaborative Linguistic And Archaeological Study). *Journal of Namibian Studies: History Politics Culture*, *34*, 3218–3240. https://doi.org/10.59670/jns.v34i.1797.

Enfield, N. J., & Comrie, B. (Reds). (2015). *Languages of Mainland Southeast Asia*. DE GRUYTER. https://doi.org/10.1515/9781501501685.

Everett, C. (2018). The similar rates of occurrence of consonants across the world's languages: A quantitative analysis of phonetically transcribed word lists. *Language Sciences*, *69*, 125–135. https://doi.org/10.1016/j.langsci.2018.07.003.

Gomes, C. (2015). Negotiating everyday life in Australia: unpacking the parallel society inhabited by Asian international students through their social networks and entertainment media use. *Journal of Youth Studies*, *18*(4), 515–536. https://doi.org/10.1080/13676261.2014.992316.

Guetterman, T. C., Chang, T., DeJonckheere, M., Basu, T., Scruggs, E., & Vydiswaran, V. V. (2018). Augmenting Qualitative Text Analysis with Natural Language Processing: Methodological Study. *Journal of Medical Internet Research*, *20*(6), e231. https://doi.org/10.2196/jmir.9702.

Ha, H. T. (2022). Lexical Profile of Newspapers Revisited: A Corpus-Based Analysis. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.800983.

Hirota, I., & Tsuji, T. (2021). Ethno-bamboo Approaches to the Pala'wan in the Mangrove Area of Palawan Island, the Philippines. *Southeastern Philippines Journal of Research and Development*, *26*(2), 47–62. https://doi.org/10.53899/spjrd.v26i2.153.

Jablonkai, R. R. (2020). Leveraging Professional Wordlists for Productive Vocabulary Knowledge. *ESP Today*, *8*(1), 2–24. https://doi.org/10.18485/esptoday.2020.8.1.1.

Juang, L. P., Simpson, J. A., Lee, R. M., Rothman, A. J., Titzmann, P. F., Schachner, M. K., Korn, L., Heinemeier,

D., & Betsch, C. (2018). Using attachment and relational perspectives to understand adaptation and resilience among immigrant and refugee youth. *American Psychologist*, *73*(6), 797–811. https://doi.org/10.1037/amp0000286.

King, K. A., & Mackey, A. (2016). Research Methodology in Second Language Studies: Trends, Concerns, and New Directions. *The Modern Language Journal*, *100*(S1), 209–227. https://doi.org/10.1111/modl.12309.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203. https://doi.org/10.1037/a0038695.

Kutas, M., Lindamood, T. E., & Hillyard, S. A. (2019). Word Expectancy and Event-Related Brain Potentials During Sentence Processing. In *Preparatory States & Processes* (bll 217–237). Psychology Press. https://doi.org/10.4324/9781315792385-11.

Larsen-Freeman, D. (2018). Looking ahead: Future directions in, and future research into, second language acquisition. *Foreign Language Annals*, *51*(1), 55–72. https://doi.org/10.1111/flan.12314.

Lewandowski, N., & Jilka, M. (2019). Phonetic Convergence, Language Talent, Personality and Attention. *Frontiers in Communication*, *4*. https://doi.org/10.3389/fcomm.2019.00018.

Madrunio, M. R. (2022). Lexical and Grammatical Features of Memoranda of Agreement (MOA) on Academic Partnerships. *Journal of English and Applied Linguistics*, *1*(1). https://doi.org/10.59588/2961-3094.1004.

Maharani, P. N., Afifah, N., & Lubis, Y. (2023). The Power of Phonology: Analyzing the Impact of Sound Structure on Language. *Journal of English Language Teaching and Learning*, *4*(1), 48–52. https://doi.org/10.33365/jeltl.v4i1.3305.

Mahmud, T., Ptaszynski, M., & Masui, F. (2023). Automatic Vulgar Word Extraction Method with Application to Vulgar Remark Detection in Chittagonian Dialect of Bangla. *Applied Sciences*, *13*(21), 11875. https://doi.org/10.3390/app132111875.

Malmasi, S., & Dras, M. (2017). Multilingual native language identification. *Natural Language Engineering*, *23*(2), 163–215. https://doi.org/10.1017/S1351324915000406.

Melinger, A. (2018). Distinguishing languages from dialects: A litmus test using the picture-word interference task. *Cognition*, *172*, 73–88. https://doi.org/10.1016/j.cognition.2017.12.006.

Menn, L., & Stoel-Gammon, C. (2017). Phonological Development. In *The Handbook of Child Language* (bll 335–360). Wiley. https://doi.org/10.1111/b.9780631203124.1996.00014.x.

Nausa, R. (2019). Modality and Code Glosses to Transition from Academic Written to Oral Discourses. *Íkala*, *24*(1), 51–67. https://doi.org/10.17533/udea.ikala.v24n01a02.

Pathan, H., Memon, R. A., Memon, S., Shah, S. W. A., & Magsi, A. (2018). Academic Vocabulary Use in Doctoral Theses: A Corpus-Based Lexical Analysis of Academic Word List (AWL) in Major Scientific Disciplinary Groups. *International Journal of English Linguistics*, *8*(4), 282. https://doi.org/10.5539/ijel.v8n4p282.

Safari, I. (2018). A Corpus-Based Contrastive Study of Code Glosses Used in English Academic Articles Written by Authors of Politics and Applied Linguistics. *International Journal of Linguistics*, *10*(2), 39. https://doi.org/10.5296/ijl.v10i2.12242.

Saimon, A., Abdul Latif, N. A., Rahim, N. A., Yusoff, A. I. M., & Mohammad, N. S. H. (2021). Students Comprehension To The New Terminology of Covid-19. *International Journal of Social Sciences and Economic Review*, 12–22. https://doi.org/10.36923/ijsser.v3i1.99.

San Jose, D. B. S., & A. Pilar, J. G. (2023). Phonological idiosyncrasies of the Southern Sorsogon dialect in Bulan, Philippines. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, *15*(3). https://doi.org/10.21659/rupkatha.v15n3.05.

Sherif, S. M., Alamoodi, A. H., Albahri, O. S., Garfan, S., Albahri, A. S., Deveci, M., Baker, M. R., & Kou, G. (2023). Lexicon annotation in sentiment analysis for dialectal Arabic: Systematic review of current trends and future directions. *Information Processing & Management*, *60*(5), 103449. https://doi.org/10.1016/j.ipm.2023.103449.

Smith, J. A. (2017). Textual Analysis. In *The International Encyclopedia of Communication Research Methods* (bll 1–7). Wiley. https://doi.org/10.1002/9781118901731.iecrm0248.

Sohayle, M. H. A. R., Gubat-Manabilang, M. G.-M., S. Rafal, G., & S. Magomnang, A. (2022). Who is Copying Who? Similarities of Words between Bahasa Melayu and Maranao: An Ethnographic Study. *Journal of Communication and Cultural Trends*, *4*(1). https://doi.org/10.32350/jcct.41.01.

Symaco, L. P. (2017). Education, language policy and language use in the Philippines. *Language Problems and Language Planning*, *41*(1), 87–102. https://doi.org/10.1075/lplp.41.1.05sym.

Tamano, R. G., Otara, F. A., & Umpa, J. M. (2021). Languages Spoken and its Influence to the Lexicons and Language Perspectives of ESL Students. *International Journal of Linguistics, Literature and*

Translation, *4*(5), 83–89. https://doi.org/10.32996/ijllt.2021.4.5.7.

Tondo, F. H. (2020). The People Of Kao And Their Language In The Northern Coast Of Halmahera. *Jurnal Masyarakat dan Budaya*, *22*(2), 75–85. https://doi.org/10.14203/jmb.v22i2.1074.

Zorc, R. D. (1974). Towards a Definitive Philippine Wordlist: The Qualitative Use of Vocabulary in Identifying and Classifying Languages. *Oceanic Linguistics*, *13*(1/2), 409. https://doi.org/10.2307/3622750.