

Legacy Language Materials in the Ernesto Constantino Collection: Challenges and Lessons for Building a Philippine Language Archive

Elsie Marie T. Or and Dustin Matthew O. Estrellado

Abstract

The late Dr. Ernesto Constantino was a prominent Filipino linguist who conducted many comparative studies on the morphology and syntax of various Philippine indigenous languages. As such, his personal collection, which was donated to the University Archives of the University of the Philippines Diliman, contains a wealth of raw data collected from different communities all over the Philippines. To date, we have cataloged 458 magnetic reel tapes and 733 cassette tapes from various linguistic fieldwork conducted by Dr. Constantino and his team of research assistants between the 1960s and the 1990s. We have also so far cata-

logged word and sentence lists in over a hundred Philippine language varieties. Many of these languages and dialects are under-studied while some, such as Yogad, Bolinao, and Manide, are also currently marked as threatened or endangered languages. This paper presents a preliminary assessment of the state of the legacy language materials in the Constantino Collection. We identify challenges that we face in cataloging and archiving the materials and the issues that should be faced in attempting to get subsequent use of the materials.

Keywords: language archives, language documentation, language preservation, Philippine languages and dialects, endangered languages

1 Introduction

The late Dr. Ernesto Constantino (1930–2016) was a prominent Filipino linguist who took an active role in the adoption and development of Filipino as the Philippines’ national language. Constantino is considered as one of the pillars of linguistics in the Philippines and his research primarily focused on the comparative study of the morphosyntactic structures of various Philippine indigenous languages. Among his notable works are “The Sentence Patterns of Twenty-six Philippine Languages” (1965), “The Deep Structures of Philippine Languages” (1970), “Tagalog and Other Major Languages in the Philippines” (1971), and “The ‘Universal Approach’ and the National Language of the Philippines” (1974).

Legacy Language Materials in the Ernesto Constantino Collection

From the 1960s up until his retirement in 1993, Dr. Constantino led a number of big, long-term research projects through which he and his team of research assistants and graduate students at the University of the Philippines Department of Linguistics (UP Lingg) gathered language data from as many Philippine languages and dialects as possible. They mostly used a direct elicitation method, which required language consultants to translate over 2,000 test sentences and over 5,000 words and phrases into their own languages.

Figures 1 and 2 show a summary of the types of words and sentences included in the elicitation materials they developed. These images were taken from a proposal drafted on August 11, 1965 for the establishment of a Research Center for Philippine Languages which was signed by Constantino and three other members of UP Lingg at the time, namely, Ernesto H. Cubar, Marietta N. Posoncuy, and Consuelo J. Paz.

The data which they had gathered within the span of more than three decades are preserved in audio recordings and in handwritten and typewritten transcriptions. This collection of language materials was turned over by the Constantino family to the Archives Section of the University of the Philippines Diliman Main Library (University Archives) along with Dr. Constantino's other personal documents and materials, such as books, correspondences, field notes, and unpublished manuscripts, upon his death in 2016.

Due to the lack of personnel at the University Archives, only a very general description of the contents of the collection was made. UP Lingg borrowed the collection to create a more detailed catalog and to digitize the legacy language materials and other relevant documents that we hoped to find in the collection. As of this writing, we are still in the

5,315 WORDS AND PHRASES IN TAGALOG WITH ENGLISH TRANSLATION
TO ELICIT WORDS AND PHRASES IN THE DIFFERENT
PHILIPPINE LANGUAGES AND DIALECTS

(1) Lexicostatistic List (Swadesh List)	200 words
(2) Comparative Word List	230 "
(3) Noun List I	1,116 "
(4) Noun List II	1,070 "
(5) Plural Noun List	72 items
(6) Cultural Terms	36 "
(7) Kinship Terms	100 words and phrases
(8) Verb List I	878 words in sentences
(9) Verb List II	614 "
(10) Verb List III	473 "
(11) Adjective List	450 words
(12) Adverb List	86 words and phrases
TOTAL	5,315

Figure 1. Summary of Types of Words and Phrases in Constantino's Elicitation Material

process of cataloging all of the materials in the collection, which contains 211 document boxes and 70 big corrugated boxes. According to the estimated measurement made by the University Archives Section, the

Legacy Language Materials in the Ernesto Constantino Collection

2,214 TEST SENTENCES IN TAGALOG, ILUKANO, SEBUANO OR BIKOL WITH ENGLISH TRANSLATION TO ELICIT THE DIFFERENT MORPHOSYNTACTIC PATTERNS OF ALL PHILIPPINE LANGUAGES AND DIALECTS

(1) Basic Sentence List I	1,269 Sentences
(2) Basic Sentence List II	400 "
(3) Sentence Addenda I	26 "
(4) Sentence Addenda II	22 "
(5) Sentence Addenda III	50 "
(6) Addenda for Pronouns I	90 "
(7) Addenda for Pronouns II	58 "
(8) Addenda for Particles I	58 "
(9) Addenda for Particles II	45 "
(10) Conversational Expressions	196 items
TOTAL	<u>2,214</u>

Figure 2. Summary of Types of Sentences in Constantino's Elicitation Material

collection contains 34.64 linear meters of archival materials. However, this measurement was given to us before they later on discovered additional boxes in the library's storage that were apparently also part of

the collection donated by the family to the University, therefore this measurement will be updated once the cataloging and rehousing of the materials are completed.



Figure 3. The Ernesto Constantino Collection at Its Temporary Storage Area

Also not included in the above-stated count are several more boxes of over 730 cassette and microcassette tapes, which had been directly turned over to UP LINGG in 2013. The digitization of this collection of cassette and microcassette tapes is also currently ongoing under a separate digitization project headed by Mr. Michael Manahan.

We have yet to find a complete index which lists all of the language materials that are available in the entire collection—that is, if one had

Legacy Language Materials in the Ernesto Constantino Collection



Figure 4. Some of the Cassette and Microcassette Tapes in the Collection

ever been made. We have so far identified 458 magnetic reel tapes, as well as several reams of word lists and sentence lists in over a hundred Philippine languages. Some of these are available in multiple dialectal

varieties that were translated by speakers hailing from different parts of the Philippines (see §6).

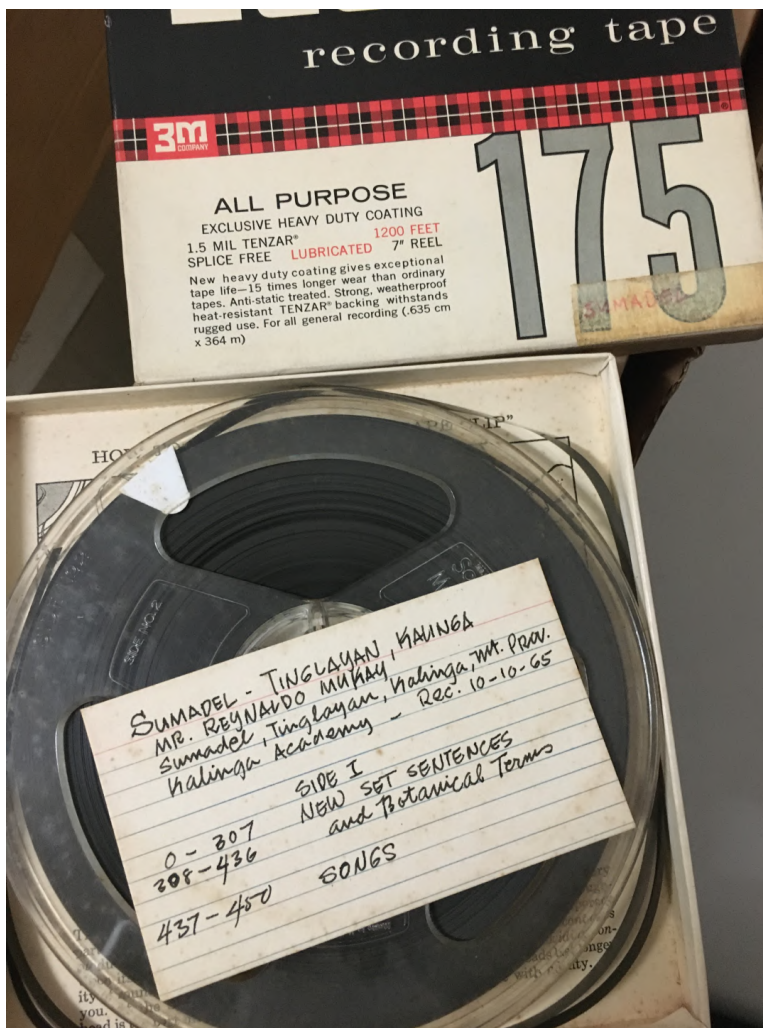


Figure 5. A Magnetic Reel-to-reel Tape From the Collection

The earliest audio recordings that we have found in the collection donated to the University Archives were recorded in open magnetic reel

Legacy Language Materials in the Ernesto Constantino Collection

tapes in 1960. Many of the cassette and microcassette tapes are undated so it is unclear when some of the recordings were created. Among the cassette tape recordings that indicate dates of recording, we found that the earliest ones were created in 1975.

The collection moreover contains multiple drafts of Constantino's ambitious final but unfinished project titled "The Universal Dictionary of Philippine Languages" (see Figure 6). This project was funded by the Toyota Foundation from 1986 to 1992 and Constantino planned for it to contain over 20,000 entries with headwords in English translated to their equivalent or closest equivalent expressions in at least 105 Philippine languages (Constantino, 1994).

Aside from these legacy language materials, the collection also includes books and journals that were previously owned by Constantino, as well as letters, fieldwork notes and reports, administrative documents, lecture notes, etc.

Constantino was also a folklorist, who was one of the co-founders of the University of the Philippines' Folklorist Society. As part of his folklore research, he had also recorded several oral narratives and folk songs from native speakers of different indigenous languages in the Philippines. A few of these had been published as part of the Endangered Languages of the Pacific Rim project of the Tokyo University of Foreign Studies in Japan (Constantino, 2001, 2002, 2003).

There are also manuscripts and materials that were created and owned by Cecilio Lopez (1898–1979) that somehow found their way into Constantino's collection. Lopez is known as the first trained Filipino linguist. He remains a significant figure in Philippine linguistics, who

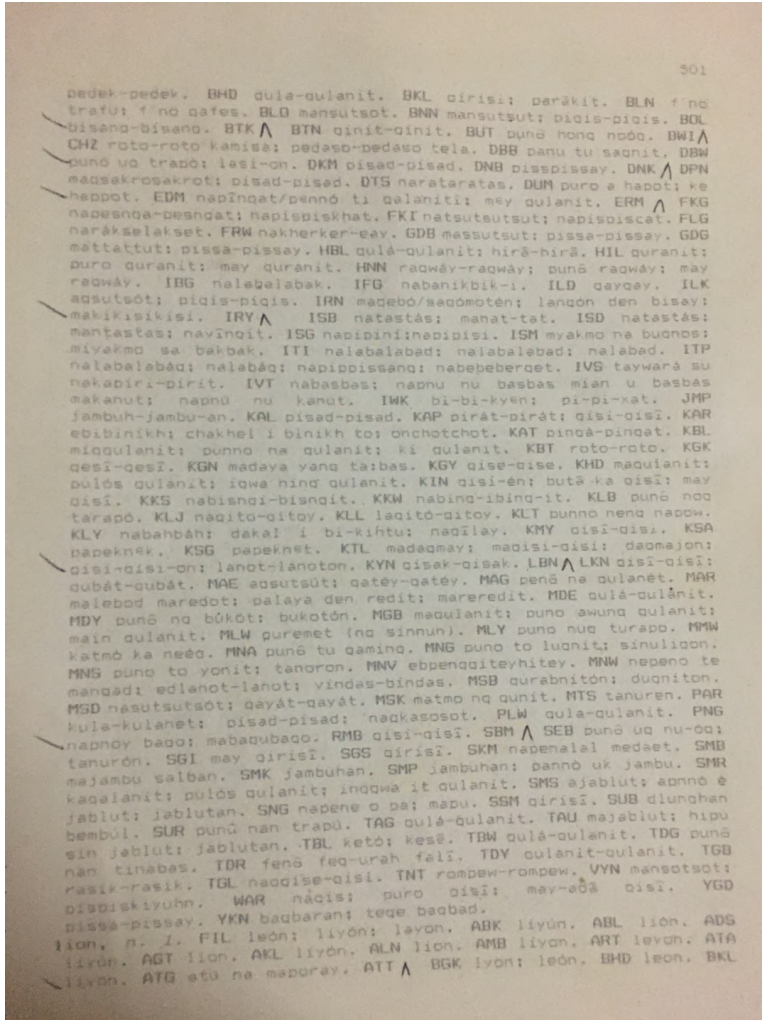


Figure 6. A Page From a Draft of “The Universal Dictionary of Philippine Languages”

exercised a lot of influence in the crafting of language policies in our country after the Second World War.

While we have yet to completely catalog personal documents that belonged to Constantino and Lopez, we believe that they might serve as valuable resources in piecing together the historical development of linguistics, of the evolving language policies in the country, as well as the development of the social science disciplines in our country, particularly in academic institutions from the 1900s to the present day. Researchers who might be interested in studying the intellectual life of the two figures might also find these academic paraphernalia of interest as well.

We should note at this point, however, that the succeeding discussion will focus on the legacy language materials in the collection that were produced by Constantino and other researchers at the UP Department of Linguistics for the purpose of language documentation and other language-related studies. The objective of the present paper is mainly to provide a preliminary assessment of these language materials and to present reflections on the prospects of and challenges in getting secondary use out of them.

2 Potentials for Secondary Use

The Constantino Collection is a veritable trove of language data from over a hundred language varieties in the Philippines. Most of the languages in the collection are still under-studied. There are also data on languages such as Yogad, Bolinao, Adasen, Bangon, Bogkalot, and Manide, which are currently tagged as threatened or endangered languages (Eberhard et al., 2022). There are also recordings and transcribed word and sentence lists in languages which are already nearly extinct,

such as Ata, as spoken in Negros Oriental, and Ratagnon, whose native speakers' ancestral domain is located in Occidental Mindoro.

There are plenty of opportunities for secondary use of these legacy language materials:

- They might serve as jumping-off points for the creation of language descriptions, reference grammars, and glossaries or dictionaries. These in turn could be used in developing materials for language revitalization projects and/or language instruction.
- The data in this collection could be used for comparative, typological, and diachronic linguistic studies.
- The data recorded in these materials might reflect earlier forms of the language and therefore could be used to trace how languages have changed over several decades and the forces that might have contributed to language change.
- These data can also be used to train linguistics students by providing them with hands-on experience in transcribing, processing, and analyzing linguistic data. This would especially be useful in circumstances where students are not able to conduct fieldwork and gather data on their own.

Aside from these possible uses of the legacy materials in linguistics-related research and language development, the audio recordings of various oral narratives and folk songs will also be valuable to the documentation and study of Philippine folklore, culture, social history, and ethnomusicology. They also have the potential of giving opportunities to the descendants of the speakers, who were recorded in the reels and tapes, as well as other members of the various ethnolinguistic commu-

nities from which the data were gathered from, to reconnect with and rediscover their familial and/or cultural heritage.

3 Challenges in Archiving and Using Legacy Language Materials in the Collection

As with other legacy language materials found in the archives of other institutions in other parts of the world, the materials in the Constantino Collection present several challenges to us who are in the process of archiving and digitizing them, and to anyone who would want to get secondary use out of them

Austin (2017) identified four broad areas where legacy materials commonly present challenges to archivists and secondary users. These include (a) the form of the materials, (b) the content of the original recorded materials, (c) the analyses of the data recorded, and (d) the context related to the recording of the data. The following subsections will discuss each issue as pertains to the legacy materials in the Constantino Collection.

3.1 Challenges With the Form of the Legacy Materials

In terms of the form of the legacy materials, as previously noted, the materials come in both written text and audio formats. The audio recordings in the collection are in three types of media: (a) magnetic reel tapes, (b) cassette tapes, and (c) microcassette tapes. The open magnetic reel tapes also come in two forms:

1. polyester-based reel tapes (an example of which is shown in Figure 5) and
2. acetate reel tapes (an example of a silicone-coated acetate reel tape is shown in Figure 19).

The two types of reel tapes are prone to different types of problems, with the former being more prone to turning sticky when exposed to high humidity (a condition called “sticky shed syndrome”), and the latter being more prone to brittleness when they lose their plastic coating (Arton, 2015). Polyester-based tapes are also more prone to becoming deformed when it is too tightly wound (Lacinak et al., 2017). The physical condition of all of the audio recordings in the collection have yet to be comprehensively assessed by professional audio preservationists.

It should be emphasized that there is a real urgency in digitizing the open magnetic reel tapes. Some of the tapes in the collection are already moldy and, according to a report issued by the National Film and Sound Archive of Australia (2017), magnetic media not digitized by 2025 will most likely be lost forever due to the possible degradation of the materials. Working machines that can be used to play the recordings are also increasingly hard to find.

Magnetic media need to be stored in an environment with limited variation in temperatures and humidity, and where dust and other particles would have little chance of getting on the tapes (Van Bogart, 1995). Table 1, adapted from Arton (2015), shows structural considerations for storing audio collections and a list of what audio materials should be protected from.

Temperature and humidity are particular concerns when it comes to open magnetic reel tapes, which is why these are mentioned in many

Table 1. Structural Considerations and Sources of Damage

Structural Considerations	Possible Sources of Damage
<ul style="list-style-type: none"> • climate control • ventilation • floor loading capacity • fire suppression systems • security 	<ul style="list-style-type: none"> • light • flooding • earthquakes and other vibrations • magnetic fields • pest infestations

care and preservations guidelines for audio materials (e.g., Arton, 2015; Audio Engineering Society, 2006; Byers, 2003; Lacinak et al., 2017; Van Bogart, 1995). However, a consensus on the best temperature and humidity levels has not yet been reached in the industry. Generally, Arton (2015) writes that for long-term storage, a temperature range of 8°C to 12°C (46°F to 53°F) and a relative humidity (RH) level between 25% to 35% have been reported as being part of common best practices. Lacinak et al. (2017) also writes that temperature should not fluctuate more than +1°C (3°F), while RH levels should not fluctuate more than +5% within a 24-hour period. Proper ventilation and a good air conditioning system are therefore required if a collection is to be preserved for a long period of time.

When it comes to digitizing the audio recordings in the tapes, finding relatively low-cost machines for digitizing the cassette tapes was not difficult, however, the resulting digitized files that these machines can produce are in MP3 formats. MP3 is not ideal for archiving as it produces compressed, lossy files with some audio data not captured and therefore audio recordings in this format is not deemed to be ideal

for acoustic analysis (Kung et al., 2020). However, since most of the audio recordings were done in less than ideal environments in the field, they likely would not have been ideal for such type of analysis at the outset. The digitization and preservation of these audio materials in the most accessible format possible by using machines that are more readily on-hand are therefore better done sooner than later to avoid completely losing the recordings.

We still have yet to come up with a solution for digitizing the hundreds of magnetic reel tapes that are in the collection. There are institutions and private companies which already have technology for digitizing reel-to-reel tapes. However, funds still have to be raised in order to avail of these services. Striking a partnership with institutions, which already have the equipment and trained technicians for preserving and digitizing audio recordings, might be ideal as they might also already have readily publicly accessible and sustainable digital repositories. We can also weigh the options of whether or not it would be better instead to procure audio digitization machines and hire or train personnel who will be in charge of preserving, digitizing, and managing the storage and access to the audio materials. This would allow us to collect and digitize more legacy materials that might turn up from other alumni or researchers who might be interested in storing their data on Philippine languages in an academic institution such as ours.

Not all of the audio recordings had been transcribed immediately after they were recorded (or if they had been, we have yet to find the transcripts). Based on the written language data available in the collection, we can surmise that Constantino prioritized the transcription of elicited words and sentence constructions. This is likely due to the

Legacy Language Materials in the Ernesto Constantino Collection

fact that, based on his published research work, he largely focused on the description of the grammatical structure of these languages and the comparative analysis of their various typological features. The objective of these types of studies was to identify similarities or what could be considered as “core features” in the grammatical systems of these related Philippine language. This was also part of Constantino’s objective of utilizing a “universal approach” in developing the Philippine’s national language (Constantino, 1972/2015). These data were also used in the construction of Constantino’s final but unfinished project, *The Universal Dictionary of Philippine Languages*.

The written language materials require sorting as there appears to be duplicates included in the files. While some data have been typewritten (see Figure 7), there are also many handwritten texts, which can be challenging to decipher, especially those that are written in cursive (see Figure 8). This is also true for the field reports we found in the collection that were written by Constantino’s research assistants. Some of these texts also contain erasures, corrections, and marginalia, which all need to be analyzed and traced to the pieces of information that they might be commenting on.

Phonetic transcriptions were found in some of the materials. These seem to employ symbols based on the International Phonetic Alphabet. While most of the Philippine languages that were documented do not have fully standardized orthographic systems, a larger portion of the written language materials nonetheless make use of the Roman alphabet. Many documents also contain markings, which will have to be interpreted as to what they were meant to denote. Underlining of certain words, for example, can be seen in Figure 9.

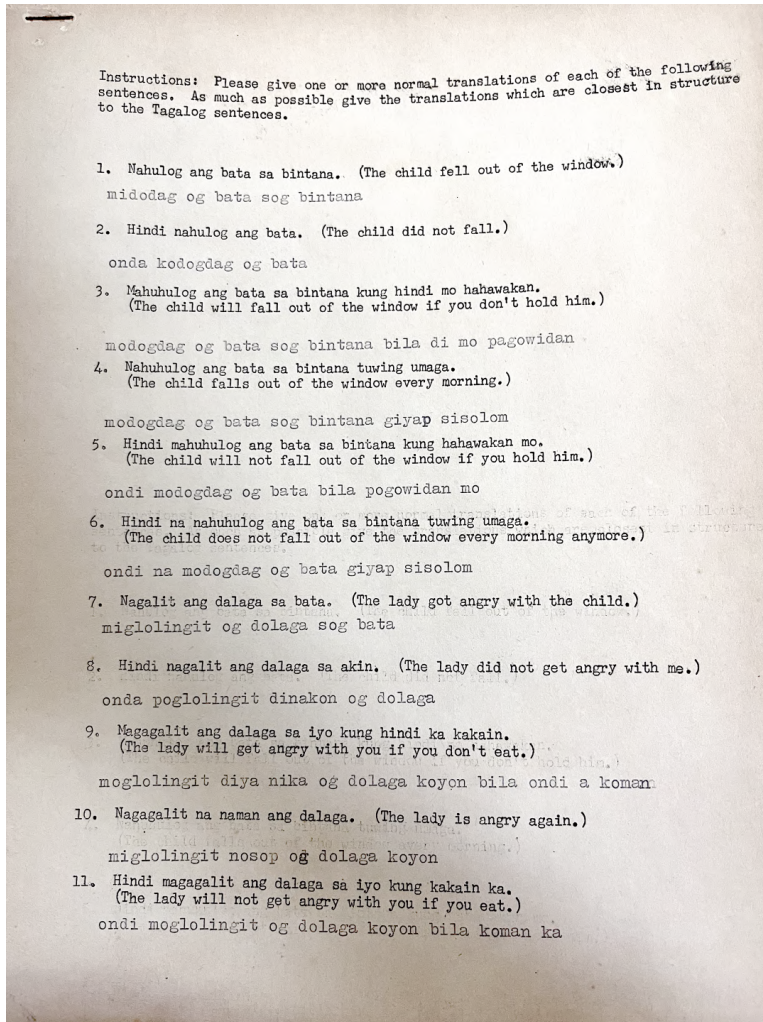


Figure 7. A Page From a Typewritten Sentence List for Subanen

3.2 Challenges With the Content of the Materials

Among the issues identified by Austin (2017) with regard to the content of legacy language materials is the appropriateness of the content that had been recorded, which might be considered taboo or inappropriate

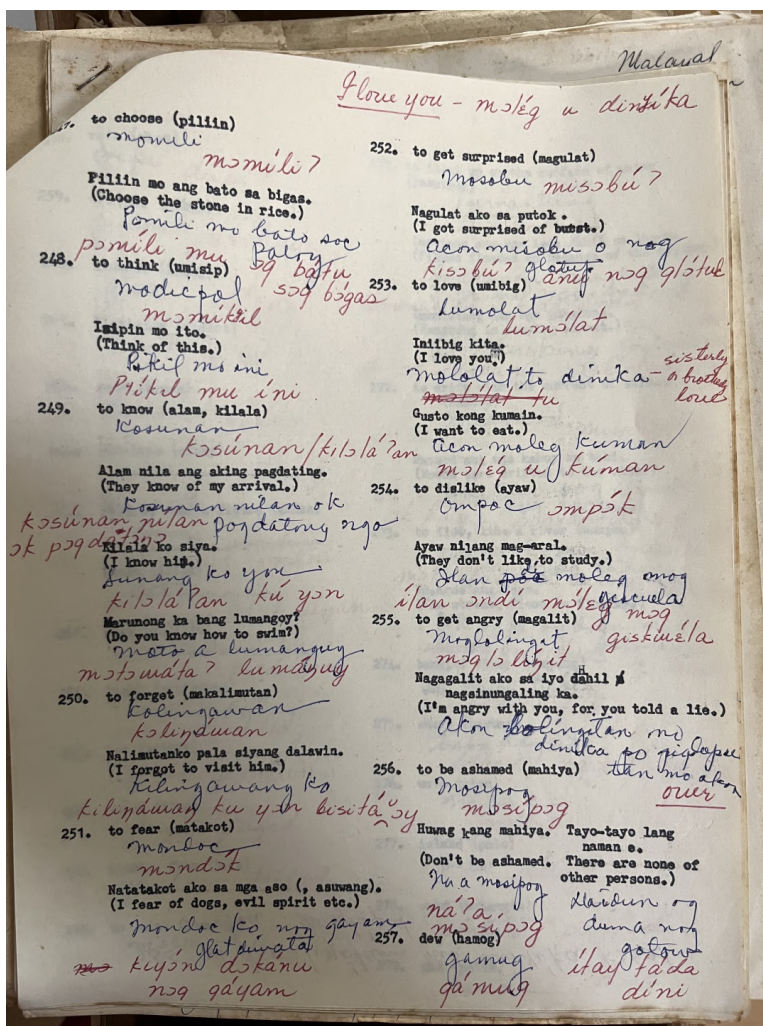


Figure 8. A Page From a Word and Phrase List With Subanen Translations and Broad Phonetic Transcriptions

according to current standards of acceptability and sensitivity. It is also possible that the original documenters recorded data that were not

J-3

1.	mabalahibo	<u>balhituan</u>	hairy, feathery <u>bambulon</u> <u>binumbol</u>
2.	mabalasik	<u>bangis/isug</u>	fierce, ferocious <u>mulinin</u> <u>boluk</u>
3.	balasubas	<u>kuriput</u> , <u>balasubas</u>	tightwad, one given to bargaining, not punctual in fulfilling obligations or debts <u>mutikas</u> , <u>dik mononan</u> <u>nsk pasad</u>
4.	balingkinitan	<u>niwang</u>	slim, slender (said of the body) <u>mula kpis</u> <u>molunggas</u>
5.	baliktad, baligtad	<u>nabali</u>	inside out, upside down <u>mikulatik</u> <u>misawik</u>
6.	baliw	<u>buang</u>	crazy - <u>migborag</u> , <u>binorag</u> <u>boragon</u>
7.	baluktot	<u>balikug</u>	curved, bent, twisted <u>mipokok</u> , <u>dunkal</u> , <u>buktat</u> <u>migbukudul</u> , <u>mipilas</u> , <u>glingig</u>
8.	banal	<u>santus</u>	pious, religious, virtuous <u>palotanud</u> , <u>miktanud</u> , <u>midulus</u>
9.	bantog	<u>inila</u>	popular, famous, illustrious <u>mibantog</u> , <u>Rusuligan</u> , <u>Kusungundulan</u> , <u>Mukudunding</u> <u>Musubuh</u>
10.	mabantot	<u>lasu</u>	stinking, fetid (usually said of liquids) <u>mumukat</u>
11.	mabangis	<u>bangis/isug</u>	ferocious, wild, cruel <u>mulentin</u> , <u>mila</u> , <u>bulok</u> <u>lulingiton</u>
12.	bangas	<u>bungi</u>	harelip, wounded on front part of face <u>sungeh</u> , <u>misungeh</u>

Figure 9. A Page From a Subanen Word List Where Underlined Portions Can Be Seen

meant for wide distribution, either due to privacy concerns or to protect information related to indigenous knowledge and practices.

Legacy Language Materials in the Ernesto Constantino Collection

This issue is not very prevalent in the written language data in the collection, given that they mostly involve translations of context-free words and sentences. However, personal information included in the elicited materials (e.g., full names and addresses of language consultants) and in the field notes of researchers would have to be treated carefully.

Audio recordings, on the other hand, need to be replayed and checked for content that might be of more sensitive nature or which might not have been meant for public access. For example, there are narrative accounts of killing people and altercations between two groups. There are also recordings of ritual songs and chants, for which the UP Lingg might require permission from the ethnolinguistic groups wherein these traditions are passed down and ask if these can be made available publicly or for limited access only according to a specific set of guidelines.

Below is an example of a listing found on the labels included inside the box of one of the reel tapes. The symbols, capitalizations, and indentations originally used in the label are preserved as much as possible in the transcription below.

“ILONGOT — Binuangan, Belansi, Dupax, NV.

Rec. 2/29 [sic]/67

Singing –

SIDE I

0–21 IBALOY Gospel Song

23–37 ILONGOT COURTSHIP SONG (Kámma)

PISANO ANGGAMAK

40–52 LINUP LINUP (song for the Dead)

ILONGOT — TEGEP, BELANSI, DUPAX, N.V.

55–274 LX Ma. Ignacio “Yapugo” Lukba” ”

SIDE II — ILONGOT /Binuangan, Dupax

March 1, 1967

- 0–19 Angagáked – song, offering to spirits after hunt by Syengba
(the hunter’s name)
- 20–32 “Nayanud tamlubat...” song by “Yapugo”
(The Lost Ring in River) Mr. Ignacio Lukba
- 35–57 Story “Turtle & The Monkey”
- 59–70 “The Hunting Story” by Syëngbë
- 72–84 “The Hunting Story” by Kamma Utaw
- 86–121 Historical Account 18 Japs Killed by 23 Ilongots
- 123–130 Dëmo (kaingin song)
- 141–145 Farmers’ Prayer (song by Mrs. Mago Pulatingat)
- 145–158 Happiness of farmers (song by Mrs. Mago Pulatingat)
- 159–172 Linup Linup song by Kamma Utaw
- 175–202 TAGAPÁNDET (Victor’s song) to the spirit of the defeated
dead by Kamma Utaw
- 204–211 Dumyeke – song about eating in mt.
- 212–265 LÍT LIT (Native violin) by Taddem
- 267–295 SONG by Taddem
- 297–311 Dëmo (Song to his task to be finished)
- 312–336 Speech by Taddem
- 338–359 Morning Bath Song – Kamma Lutaw
- 359–364 (Explanation of song) – Kamma Lutaw
- 364–405 “Ana’na” Welcome Song – Kamma Lutaw
- 407–471 Competition Song (in Love)
- To end “Nga dey Pakitumbeg’an moy at tam degen muy”

[Marked “2”]

The provided titles for the content in the audio recordings are usually broad or general and do not indicate specific details about the recorded

songs or narratives. The sample transcription above provides some specific details for some content (e.g., number of Japanese soldiers killed by how many Ilongots; what the song titled “Dumyeke”), but some do not (e.g., we do not know what the speech by Taddem is about; nor what the song title “Nayanud tamlubat” means and what it is about). The names of the people who were included in the label of some content (e.g., Kamma Lutaw, Taddem, Mrs. Mago Pulingat), while other were not. Collaboration with speakers of the languages recorded in the tapes will have to be sought in order to re-check, transcribe, and decide whether or not the contents of the audio recordings could be made available to the public or at least to future researchers.

3.3 Challenges With the Analyses of the Data Recorded

The data included in the collection would be generally considered as raw data as they are not accompanied by morphological parsing and glossing. Figure 10, for example, shows a short passage which is surmised to be a story about the founding of the municipality of Bolinao, typewritten on a half sheet of paper. The title is handwritten in cursive and reads “last story – Founding of Bolinao,” under which the date June 27, 1966 is written. We have yet to find the translation of the story and information about who narrated the story, who recorded and transcribed the story, how the story was collected, and for what research project the story was collected. We know that multiple people (Constantino’s team of students and research assistants) conducted the fieldwork and worked with language consultants in order to collect the data that comprise the

collection, therefore we can see different handwriting styles, note-taking styles, and labelling methods throughout the collection.

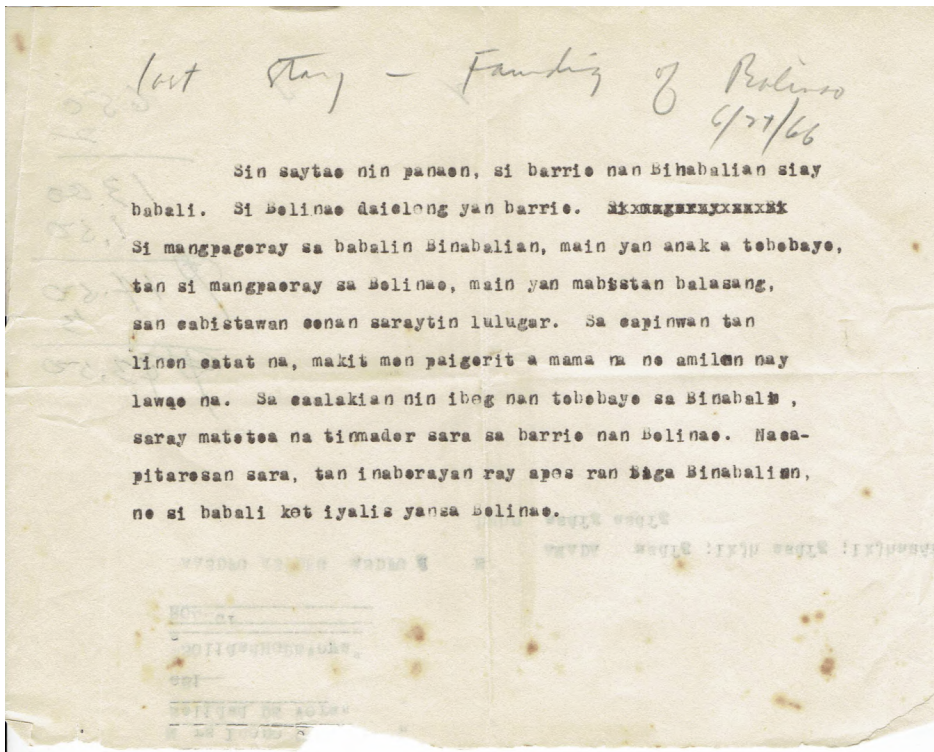


Figure 10. A Story About the Founding of Bolinao

In Figure 11, the motivation behind glossing only certain selected words in the Sambal text is unclear. However, this might be an early attempt of the researcher at analyzing the text based on the accompanying English translation. A completed clean transcription with glossing of this text has not yet been found in the collection.

The data in the collection do not include any indication as to whether or not they had been used in any completed study, or if they had

Yay Pinangibatan nin Pangalan
Palawig

before first years
 Hin unay pandur main nakin luway kastila
 in one
 ha asay lugar. Ka la tandâ nil anyay pangalan
 the place they went to
 nin lugar a kinâ la. Natakit kilan asay lalakin
 main awit nin lalawig. Dinumanî kilâ kinâ
 rope to lead came near
 Ta pinastang la'y lalaki nil anyay lugar yadtaw.
 Halitâ
 Kastila a pamastang la kayâ kay narentindyan
 here thought if asked
 nin lalaki. Yay wana nil ampastangin kay
 nang e coming that it answered
 ngalan nin awit na kayâ imbalita na
 kundan, lalawig, lalawig yadtâ awit ku.
 what he heard the response
 yay lungû lan kilay luway kastila kint
 Palawig kayâ paibat hin yadtaw pinanga-
 place
 lanan la yay lugar nin Palawig.

Palawig

Figure 11. Origin of the Name “Palawig” In Sambal

been published in some platform and in whatever format. If grammar sketches or reference grammars were to be made using the data in the collection, some effort will have to be exerted to trace these and organize cross-indices.

The Origin of the name Palawig
Rec. 12-11-65 (Palawig)
Once upon a time there came two
Spaniards, to a place. They did not know
the name of the place they were in.
They saw a man who had a big rope
They approached him & asked the man
the name of the place. They asked him
in Spanish, that's why he didn't
understand. He thought they were asking
him what he was carrying that's why
he answered "rope (a) rope is what I'm
carrying". What the Spaniards heard was
Palawig, & so from that time on
they named the place Palawig.

Figure 12. Origin of the Name "Palawig" In English

Phonetic transcriptions of the data are available in the collection for some of the collected word lists and sentence lists. However, cross-indexing of transcripts and their corresponding audio recordings is not completely systematic.

Legacy Language Materials in the Ernesto Constantino Collection

The way that transcripts are cross-indexed with their corresponding audio is via the name of the language consultant who was recorded. Figure 13, for example, includes the name (“A. Arbison”) and location (“Luuk, Sulu”) of the Tausug language consultant, handwritten on the upper right corner. Only the first letter of the person’s given name is provided, but thankfully, only one Tausug speaker with the same surname was found to have audio recordings in the collection.

In the handwritten translations of a sentence list from a Yogad language consultant shown in Figure 14, there are what seem to be accent markings over certain syllables written in blue ink. However, information is lacking about the source of the data and, while there is a label which reads “recorded” and what is presumed to be the date of the recording, it is unclear whether this means that an audio recording of the language consultant was made or if it refers to the date when the written translations were made. As we have not found a Yogad audio recording in the collection, it is also not yet possible to verify the accent or stress markings placed by the unknown annotator.

Aside from these, we also find rough notes on the lexicon or grammar of certain languages. However, these often lack explicit explanations and some labels or abbreviations used in the notes need to be interpreted by secondary users. Figure 15 is a page showing what seem to be morphological paradigms of certain root words in Bolinao. Whether or not the words written down are indeed Bolinao need to be verified because, while it was found in an envelope labelled as “Bolinao” and together with other word and sentence lists in Bolinao, the label on top of the page reads as “Bol” with a question mark in parentheses. Abbreviations were also used, as we can see in the first set of words:

Accounted by ^{3/25} *g = backed g*
ng = h

Tausug (Lulu, Sulu)
 Mr. A. Arbisson

N-1

a abaká (abaca)	abay (companion, best man)
[<i>la=mit</i>]	pandalá
b abó (ashes)	agahan (breakfast)
[<i>ábu</i>]	kakagún mahinaqát
c agimat (amulet)	agiw (cobweb, soot)
[<i>hámpan</i>]	bary lawáq (cobweb)
d akdá (literary composition)	aklat (book)
(no term)	buk
e alaga (pet)	ahas (snake)
[<i>ipátán</i>]	hats
f alaala (remembrance, souvenir)	alak (wine)
[<i>panuntumán</i>]	álak
g alakdan (scorpion)	alamang (tiny shrimp)
[<i>kádjalangkin</i>]	áyap
h alamat (tradition, folklore)	alapaap (clouds)
[<i>katákráta</i>]	gáun
i ali (aunt)	aligi (ovary of crustaceans)
[<i>inaqún</i>]	tálu
j alikabok (dust)	allá (servant)
bagúnbun	daraqakún
k alimango (a species of crab)	alimang (a species of crab)
kagáng	kagáng kagáng
l alimpuyo (whirl, eddy)	alimom (vapor rising from ground)
bue-lihúq	(no term)
m alipato (flying ember)	alipin (slave)
káyu maglupád	banyáq
n alipunga (athlete's foot)	elmirés (small stone mortar)
hawás	anáq pipisán
o alon (ocean waves)	aluloi (rain pipe, rain gut)
alún	hawúngan
p alupihan (centipede)	amá (father)
lahípan <i>lahípan</i>	ámaq
q amag (mold, mildew)	ampalaya (bitter melon)
hawásq	pa-lyáq = <i>pa-lyáq</i>
r ampaw (ball of sugared popcorn)	anahaw (palm tree)
gagatfq	nípaq

Figure 13. First Page of the Broad Phonetic Transcription of a Tausug Word List

(a) *pt.*; (b) *ft.*; (c) *p-*; and (d) *imp.* The meaning of these abbreviations will have to be deciphered and verified via consultation with native speakers.

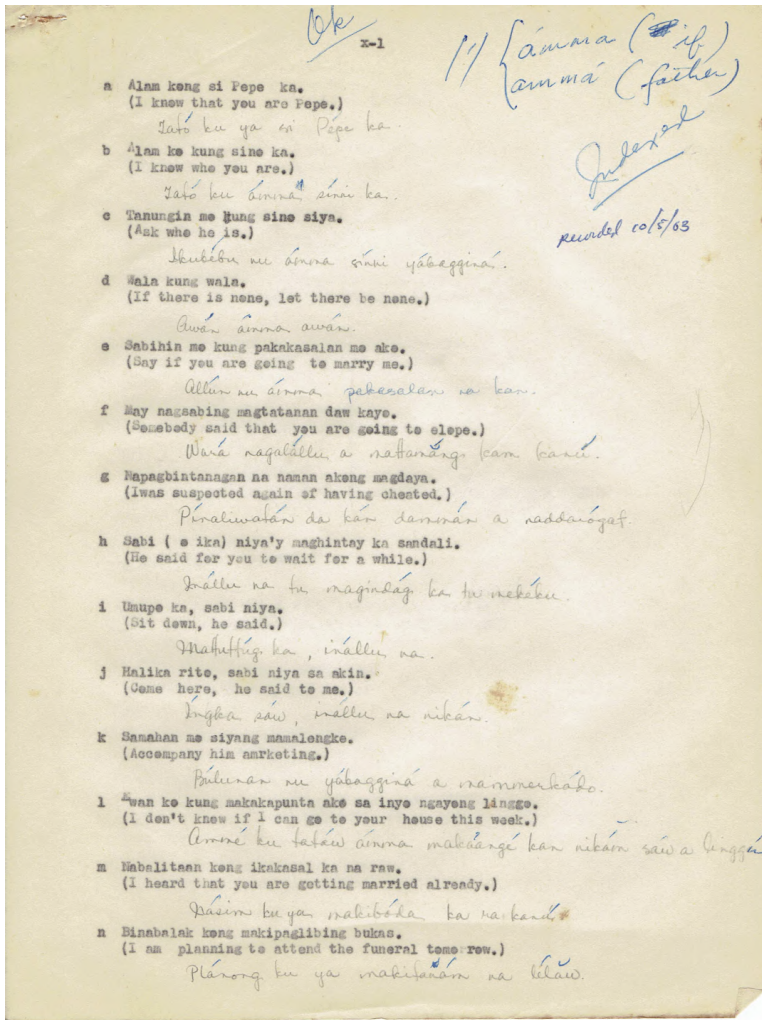


Figure 14. Accent Marking in a Yogad Sentence List

3.4 Challenges With the Lack of Context

As Austin (2017) writes, “some of the most difficult issues to deal with in legacy text materials relate to the lack of metadata (data about the data) and the meta-documentation (information about the context of

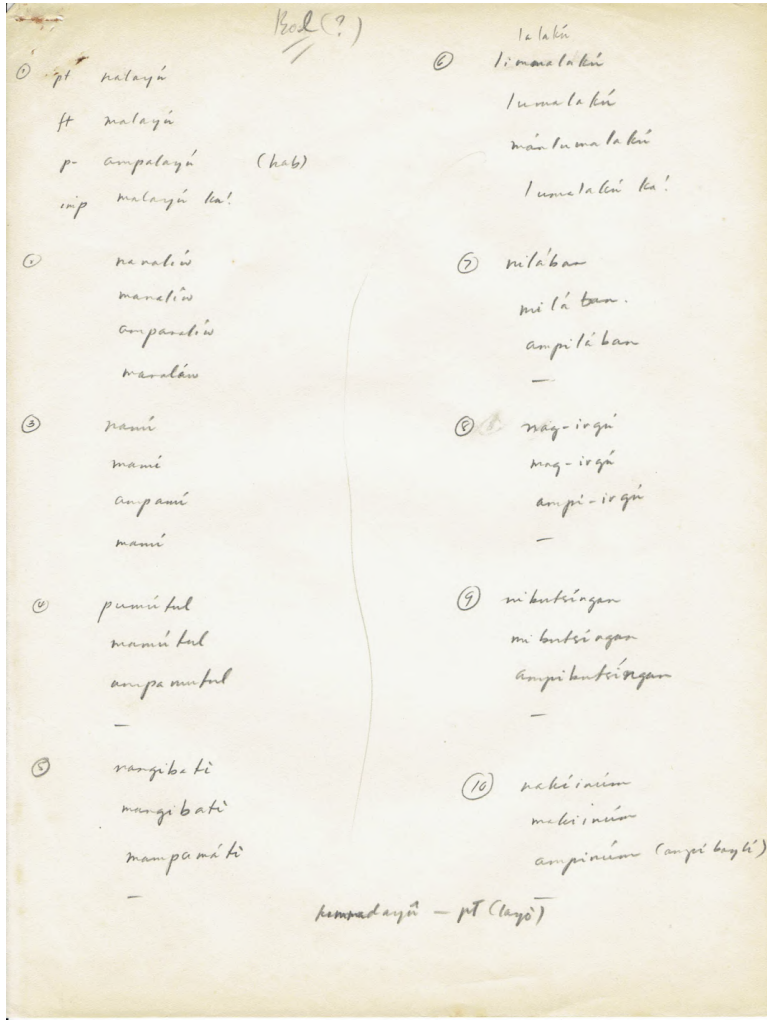


Figure 15. Notes on What Might Be Bolinao Sample Word Paradigms

collection and analysis)” (p. 37). Information about the speakers who provided the data could be missing. Austin also asserts that information about the collector of the data could also be valuable and particularly useful in interpreting how the data was analyzed, annotated, and tran-

Legacy Language Materials in the Ernesto Constantino Collection

scribed, and understanding the nature and the context surrounding the legacy language materials. This type of data could include the researcher's background, the languages that they have knowledge of, their research training, as well as details about the objectives of the original project for which the data were collected. The legacy language materials in the Constantino Collection are lacking in metadata related to the language speakers recorded and to the researchers who created the recordings, transcriptions, and annotations on the data.

Because the data were collected over a span of decades and by multiple researchers, the labelling and annotation formats also vary as well as the amount of background information that were noted down particularly on the written language materials.

LEXICOSTATISTIC LIST FOR PHILIPPINE LANGUAGES (Revised 1967)
Ernesto Constantino

ESSENTIAL DATA

Language or dialect Sorsoganon

Native speaker(s) Emil - A. Eco

Sitio, barrio, town or city, province Sorsogon Sorsogon (Laneta - Bacon)

Transcribed _____ Recorded 9/8/72
BEP

Figure 16. “Essential Data” About a Sorsoganon Language Consultant

The word and sentence lists in the collection usually only contained the following fields: (a) name of the language or dialect, (b) the full name of the native speaker, (c) their hometown or city, (d) the name of the transcriber, and (e) the date of recording. Figure 16 is an example of the form where these data are noted down. We can see that, instead of filling in the “transcribed” field, the researcher signed their initials under the date of recording. There is an additional note stating where the language consultant’s hometown is. However, it is unclear whether the field “sitio, barrio, town or city, province” is the place where the informant grew up in, where they are currently residing, or where they answered the elicitation material. Other information about the language consultant that cannot be discerned but which might be relevant include their history, age at the time of recording, details about their place of residence, and other languages that they might also be fluent in.

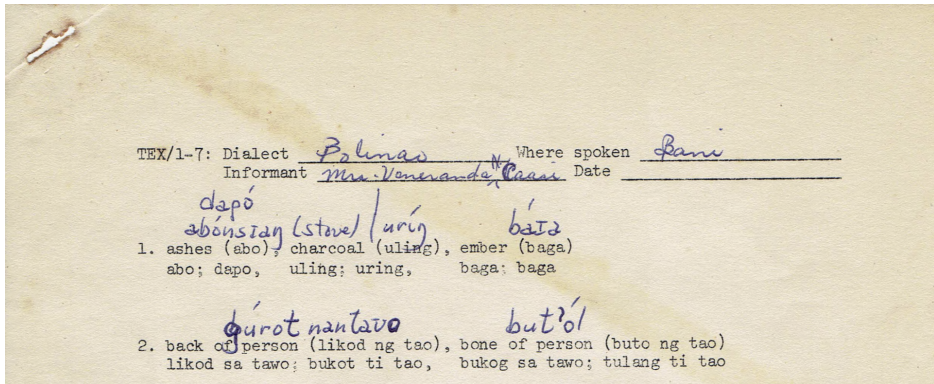


Figure 17. Limited Background Information on a Bolinao Word List

Some of the materials in the collection have even less information as we can see on the first page of the word list in Figure 17, where only the name of the language or dialect and the language consultant as well

as the place where the dialect is spoken were provided. The field for the date was left empty. No additional information was also provided about where the place Bani is located.¹

There is also an uneven level of informativeness in the labels of the audio recording. Most of the reel tapes include note cards describing the contents of the audio recordings, while there are a few where the description is written on the box itself. The former generally provide more information.

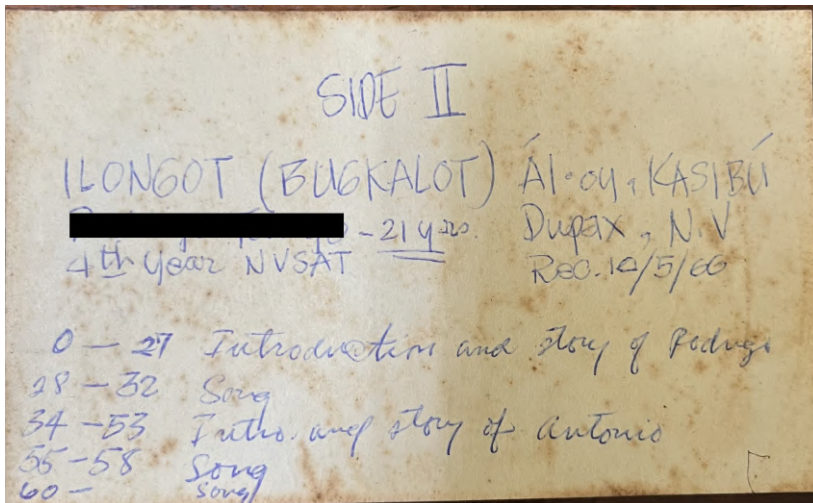


Figure 18. Notecard Describing Contents of an Ilongot (Bugkalot) Audio Recording #01

In Figure 18, we learn the name of the speaker recorded in the reel, their age, what is assumed to be their year level in college and the name of their school. It is unclear whether the recording was done in Dupax,

¹There is a municipality called Bani in the province of Pangasinan, which is among the places where Bolinao is spoken.

Nueva Vizcaya or if this is the speaker's hometown. The date of recording is also indicated.

We can also see in Figure 18 that we have less information about the second speaker recorded, who is simply referred to by his given name "Antonio." In the notecard description in Figure 19, the singers and performers who were recorded are only referred to by their gender. Meanwhile, only the first names or nicknames are provided in the notecard shown in Figure 20.

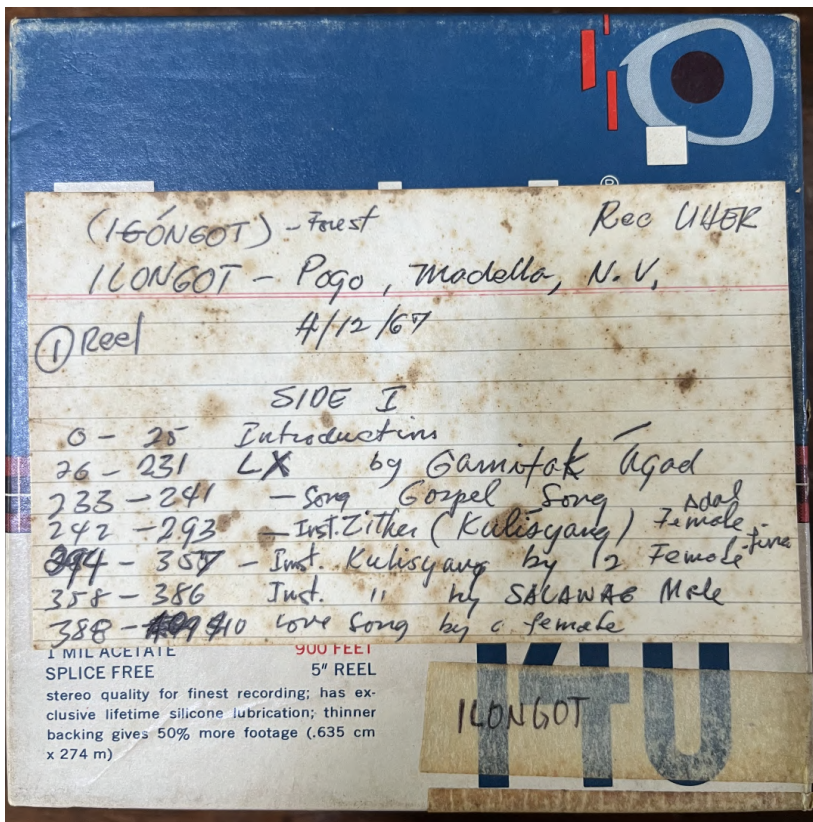


Figure 19. Notecard Describing Contents of an Ilongot (Bugkalot) Audio Recording #02



Figure 20. Notecard Describing Contents of an Ilongot (Bugkalot) Audio Recording #03

Most of the audio recording labels are handwritten, therefore there is a chance that these could be misread and copied with some errors onto the digital catalog that we are building for this project. Acronyms and abbreviations (like “LX” and “NV” as shown in Figures 19 and 20) would need to be decoded. Many are, however, usually easy enough to figure out since they appear regularly (“LX” most likely means lexicon; “NV” is a place name “Nueva Vizcaya”). Somewhat more puzzling are notes like “transferred” (to where?) or “see other reel” (which one?).

Like many of the word and sentence lists in the collection, the names of the researcher who recorded the speakers in the reel tapes are not included. The reels are not labeled with reference numbers, therefore, they cannot be easily checked against available transcripts. Overall, the organizing principle behind these legacy language materials and the Constantino Collection as a whole has yet to be figured out.

Constantino did write about the process for how they collected data for his Universal Dictionary project in a conference paper published in 1994. However, we are not sure if the procedure that he outlined there, which include how they worked together with language consultants, was something that he and his assistants were already practicing in the 1960s and 1970s as well. The description of their process and the surrounding context of the data recorded are valuable as they factor in the re-analysis and interpretation of the data, and might shed some light on the notes and corrections written by Constantino or his assistants on some of the transcripts as well.

We should note, however, that the legacy language materials in the collection were created before documentary linguistics became firmly established as its own discipline and so one could say that it is understandable why the metadata that we now recognize as being crucial to our interpretation and analyses are not present, or at least not systematically indicated and organized.

We are not yet done sifting through all of the materials in the collection though, so there is still a chance that we might find more documentation about the context surrounding the legacy language materials. Moreover, language and dialect names have to be verified. Audio recordings need to be matched to their transcripts, and fieldwork notes need

to be matched to the data collected. The repatriation and digitization of the materials, as well as the establishment of infrastructures and systems to ensure their preservation and the creation of access to or ways of transmitting the knowledge recorded in these materials all have to be planned and carried out as well. There is undoubtedly still a lot of work to be done in processing the legacy language materials in the collection.

4 The State of Language Archiving in the Philippines

While Constantino's legacy language materials essentially checks all of the issues that legacy materials pose that Austin (2017) and Dobrin and Schwartz (2021) identified, especially in their application to language documentation, description, and revitalization, we do see here a great opportunity for fulfilling the vision that Constantino and his colleagues, Ernesto Cubar, Consuelo Paz, and Marietta Posoncuy, had in the 1960s of creating a Research Center for Philippine Languages, which among other things, will house the archives of Philippine languages and dialects. Even then in the early 1960s, they had already recognized the increasingly threatened status of many Philippine languages and the need for a way to document and preserve them.

The archives that we will build, however, need to be better than the physical archives that Constantino started and the archives that might have been originally envisioned by their group. Over the decades, Constantino and UP Lingg were not able to keep up to date on and apply best practices for building and maintaining a language archive

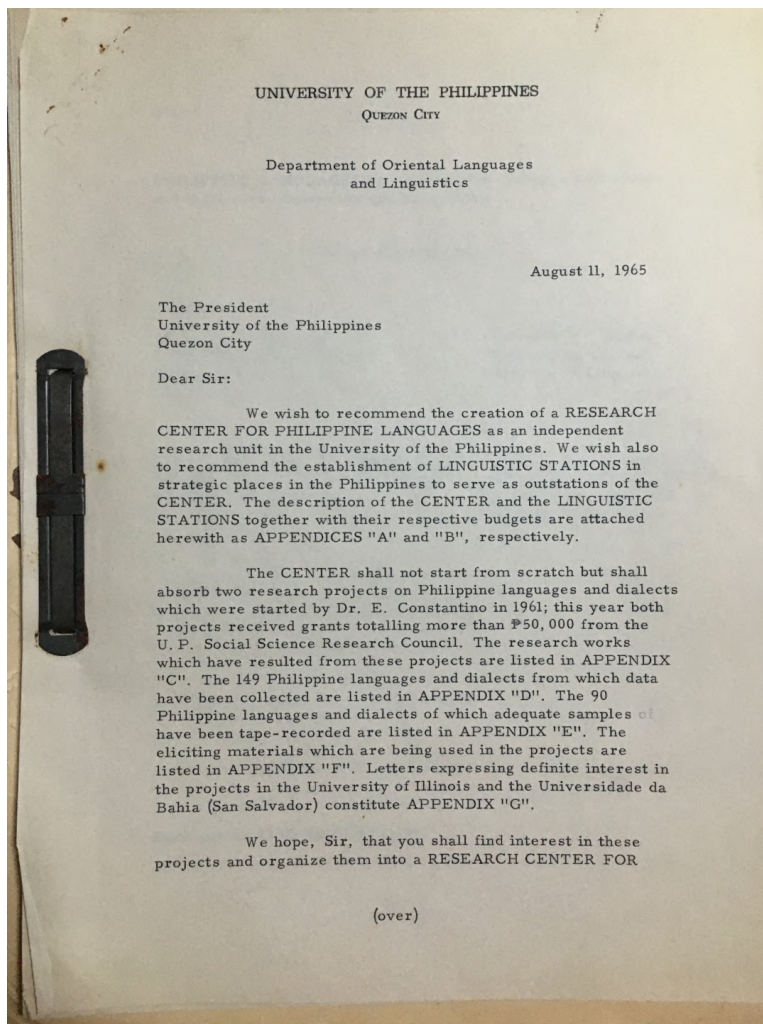


Figure 21. 1965 Proposal to Create a Research Center for Philippine Language at the University of the Philippines

that would be able to perform the following functions as described by Johnson (2004, as cited in Austin, 2011):

Legacy Language Materials in the Ernesto Constantino Collection

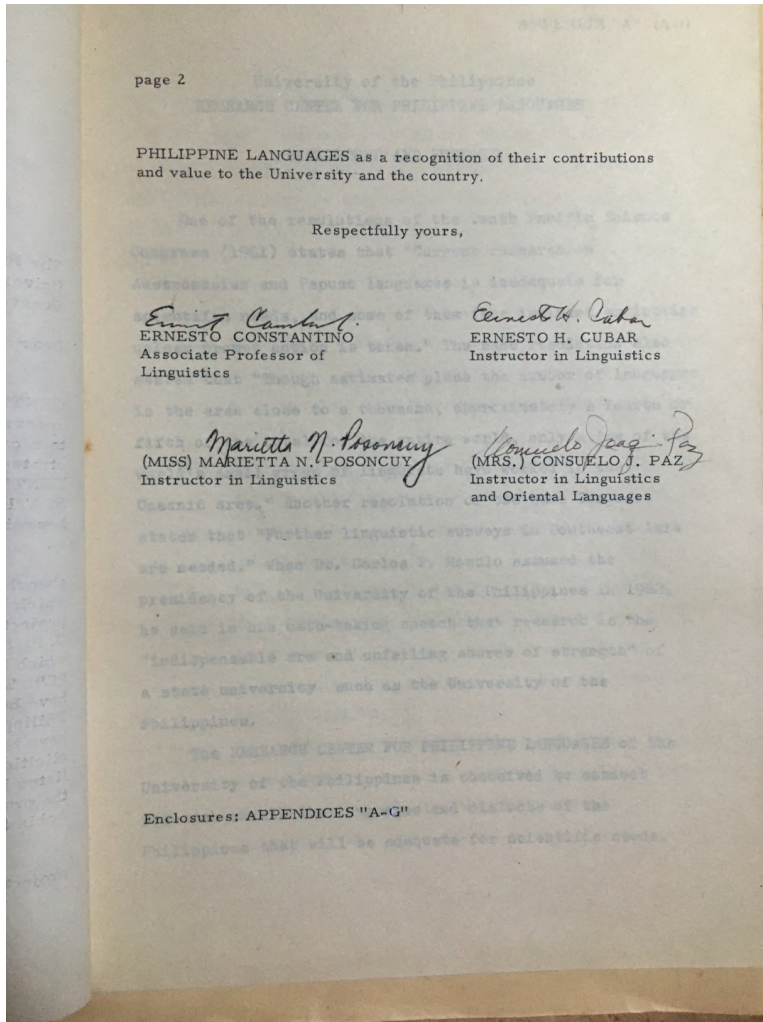


Figure 21. 1965 Proposal to Create a Research Center for Philippine Language at the University of the Philippines

1. to preserve recordings of endangered/minority languages for future generations;
2. to facilitate the re-use of materials for:

- language maintenance and revitalization programs;
 - typological, historical, comparative studies;
 - any kind of linguistic, anthropological, psychological, etc. study that other researchers might conduct;
3. to foster development of both oral and written literatures for endangered languages; and
 4. to make known what documentation currently exist for which languages.

The practice of developing and maintaining a language archive require specific technical skills, which are not usually taught as part of the curricula of most linguistics programs. Meanwhile, few people in the Philippines engage in the study of archival and preservation methods and practices due to the perceived lack of employment opportunities and value placed on the field and its practice in the Philippines.

More and more researchers in the Philippines are greatly motivated to engage in language documentation and description, thanks in part to the implementation of policies and programs such as the Mother Tongue-based Multilingual Education (MTB-MLE) Program and Indigenous Peoples Education (IPEd) Policy, which among other things require the codification and standardization of the grammatical rules and orthography of languages for them to be considered as viable mediums of instruction in the early grades. While the implementation particularly of the MTB-MLE program has been controversial and suffers from many problems, it has put a spotlight on threatened Philippine languages and the need to document, develop and/or revitalize less-studied languages.

The practice of archiving language materials collected in order to preserve them for future generations, however, has not yet caught on

due to the lack of people with skills and knowledge and the lack of funds that could provide people opportunities to work full-time on such an endeavor. There is also some hesitance on the part of researchers to make their primary data widely accessible due to ethical issues that need to be considered in order to protect the rights of indigenous people, since language is also considered part of their intangible indigenous knowledge and wealth that are protected by law. Thus being the case, we could say that language documentation as practiced by many local researchers in the Philippines has not yet been able to fully serve the mandates of the discipline as described by Henke and Berez-Kroeker (2016) and Himmelmann (1998), who write that the contemporary practice of language documentation should consider as among its top priorities the long-term storage and preservation of primary data, especially if these are endangered language materials. Still, there are, however, a few researchers who have deposited their Philippine language data in well-established digital repositories. Maria Kristina Gallego's collection of Ibatan audio and video recordings is one of the most recently archived corpus of a Philippine language documentation project on the Endangered Languages Archive (Gallego, 2019). There are also local organizations such as the Mangyan Heritage Center, which maintain their own archives and data repositories.

More and more people are becoming aware of the importance of preserving, maintaining and revitalizing endangered languages. This is perhaps as best evidenced by the United Nation's declaration of 2022 to 2032 as the Decade of Indigenous Languages. Researchers engaged in language documentation are also increasingly being made conscious of their responsibility of not only producing good descriptions of the

languages that they are documenting, but also of archiving their corpora, not only for preservation, but also for possible future use.

The issues and challenges that we have identified in the archived language materials in the Constantino Collections teach us things to avoid and things to improve. Among these is the value of creating and keeping well-organized metadata and meta-documentation of collected language data. These identified issues will also inform the creation of a more sustainable plan and system of use for language data that will be kept in what we hope will become a “living archive,” so that these data can continue to be used for the advancement of our knowledge of Philippine languages and linguistics, and also continue to benefit the communities where the data would be gathered from.

Being more active in providing and promoting open access to data and scholarship being done by our researchers is also important so that it will be easier to identify what work still needs to be done. Rather than repeating work that has already been done on certain languages, we could instead be drawing our attention to the still lesser-studied languages as well as devising other methods and questions for further enriching our knowledge about our increasingly endangered languages.

The UP Department of Linguistics is not alone in the Philippines in wanting to build a repository of Philippine language materials. The National Commission on the Filipino Language or the Komisyon sa Wikang Filipino (KWF) has recently launched its repository of Philippine languages and culture (www.kwfwikaatkultura.ph). Based on its initial content, it appears that they will be including basic information on different Philippine languages, as well as audio and video recordings of native speakers using the language. The Summer Institute

Legacy Language Materials in the Ernesto Constantino Collection

of Linguistics (SIL) Philippines has already established their Philippines Language and Culture Archives (<https://philippines.sil.org/resources>), which include their publications and the data that their members have collected from various Philippine languages since SIL Philippines was first founded in 1953.

In our case, cataloging and transferring the legacy language materials from the physical archives that form a part of the Constantino Collection to a digital archive is a first step, and there are still many questions that we have to answer and challenges to face. There is the question, for instance, of the ethics involved in making the data available more widely to the public by transferring it to a digital archive. Systems of access to the data would likely require that permissions have to be renewed or sought, so new agreements and partnerships with various indigenous cultural communities (ICCs) will need to be formed. This work alone will require a lot of resources. But the work must go on for the sake of our indigenous languages and cultural heritage. The next steps that we have to take in order to build on the Philippine language archives based on the legacy materials are also more or less clear, as guided by the lessons from what Constantino and his colleagues have already built.

5 References

Arton, Carla. (2015). Care and maintenance. In Sam Brylawski, Maya Lerman, Robin Pike, & Kathlin Smith (Eds.), *ARSC guide to audio preservation* (pp. 52–76). Association for Recorded Sound Collections; Council on Library and Information Resources; The Library of Congress.

- Audio Engineering Society. (2006). *AES information document for preservation of audio recordings: Extended term storage environment for multiple media archives*. Audio Engineering Society.
- Austin, Peter K. (2011). *Who uses digital language archives?* Endangered Languages and Cultures. <https://www.paradisec.org.au/blog/2011/04/who-uses-digital-language-archives/>
- Austin, Peter K. (2017). Language documentation and legacy text materials. *Asian and African Languages and Linguistics*, 11, 23–44. <https://doi.org/10.15026/89205>
- Byers, Fred R. (2003). *Care and handling of CDs and DVDs: A guide for librarians and archivists*. Council on Library and Information Resources; National Institute of Standards and Technology.
- Constantino, Ernesto. (1994). Problems and progress in the research project on compiling “A Universal Dictionary of Philippine Languages”: The first four years. In Bonifacio P. Sibayan & Leonard E. Newell (Eds.), *Papers from the First Asia International Lexicography Conference 1992* (pp. 149–160). Linguistic Society of the Philippines.
- Constantino, Ernesto. (2001). *Nanang: I Taguwasi anna I Innawagan* [Epic: Taguwasi and Innawagan] [An Agta Negrito epic chanted by Baket Anag]. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Constantino, Ernesto. (2002). *Ullalim Banna: A Kalinga Epic — Inyullalim Gaano Laudi, Dalupa, Pasil, Kalinga, Pilipinas* [Sung by Gaano Laudi, from Dalupa, Pasil, Kalinga, Philippines]. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Legacy Language Materials in the Ernesto Constantino Collection

- Constantino, Ernesto. (2003). *The abduction and wedding of two Negrito maidens* [Sung by Baket Anag of Maraburab, Alcala, Cagayan, Philippines]. Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Constantino, Ernesto. (2015). Ang “universal approach” at ang wikang pambansa ng Pilipinas. *Daluyan Journal ng Wikang Filipino Espesyal na Isyu Blg. 2*. (Original work published 1972)
- Dobrin, Lise, & Schwartz, Saul. (2021). The social lives of linguistic legacy materials. *Language Documentation and Description*, 21, 1–36. <https://doi.org/10.25894/lld12>
- Eberhard, David M., Simons, Gary F., & Fennig, Charles D. (2022). *Ethnologue: Languages of the world* (25th ed.). SIL International. <https://www.ethnologue.com>
- Gallego, Maria Kristina S. (2019). *Consequences of contact: Documenting Ibatan within the multilingual landscape of Babuyan Claro*. Endangered Languages Archive. <http://hdl.handle.net/2196/00-0000-0000-0014-09E8-3>
- Henke, Ryan, & Berez-Kroeker, Andrea L. (2016). A brief history of archiving in language documentation, with an annotated bibliography. *Language Documentation & Conservation*, 10, 411–457. <http://hdl.handle.net/10125/24714>
- Himmelmann, Nikolaus P. (1998). Documentary and descriptive linguistics. *Linguistics*, 36(1), 161–196. <https://doi.org/10.1515/ling.1998.36.1.161>
- Kung, Susan Smythe, Sullivant, Ryan, Pojman, Elena, & Niwagaba, Alicia. (2020). *Archiving for the future: Simple steps for archiving*

language documentation collections. University of Texas at Austin.
<https://archivingforthefuture.teachable.com/>

Lacinak, Chris, Chandler, Rebecca, Rudersdorf, Amy, & Van Mallsen, Kara. (2017). *Fundamentals of AV preservation*. Northeast Document Conservation Center. <https://www.nedcc.org/av-textbook>

National Film and Sound Archive of Australia. (2017). *Deadline 2025: Collections at risk*. National Film and Sound Archive of Australia.
<https://www.nfsa.gov.au/collection/curated/deadline-2025-0>

Van Bogart, John W. C. (1995). *Magnetic tape storage and handling: A guide for libraries and archives*. The Commission on Preservation and Access; National Media Laboratory.

6 Appendix

Working List of Philippine Language Data in the Ernesto Constantino Collection

Note: This is not yet the final list of Philippine languages that have data available in the Constantino Collection. A final list can only be supplied once the data and labels used have been verified. Many materials were labelled according to the location or hometown of the language consultants. These were cross-checked against available information on the twenty-fifth edition of the Ethnologue, and as much as possible dialects of the same language are grouped together. For example, the audio recordings labelled as “Hambali” were categorized as resources of the Sambal Botolan language. ISO labels are also supplied based on what is listed in the Ethnologue. Some labels for language materials

Legacy Language Materials in the Ernesto Constantino Collection

in the collection only indicate general names. For example, there are materials simply labelled as Bisaya or Mangyan with no other indication of which language under these subgroups the data are from.

Language	Sentence List	Word List	Audio Recording
1. Adasen (tiu)	✓	✓	✓
2. Agta, Casiguran Dumagat (dgc)	✗	✗	✓
3. Agta, Mt. Iriga (agz)	✗	✗	✓
4. Agutaynen (agn)	✓	✗	✓
5. Aklanon (akl)	✓	✗	✓
6. Alangan (alj)	✗	✗	✓
7. Alta, Southern (agy)	✗	✗	✓
8. Ata (atm)	✗	✗	✓
9. Ati (atk)	✗	✗	✓
10. Ayta	✗	✗	✓
11. Ayta, Abellen (abp)	✗	✗	✓
12. Ayta, Magbukun (ayt)	✗	✗	✓
13. Bagobo-Klata (bgi)	✓	✓	✓
14. Bantoanon (bno)	✓	✓	✓
15. Bikol (bik)	✓	✓	✓
16. Bikol, Buhi'non (ubl)	✓	✓	✓
17. Bikol, Central (bcl)	✓	✓	✓
18. Bikol, Libon (lbl)	✓	✓	✓
19. Bikol, Miraya (rbl)	✓	✓	✓
20. Bikol, Northern Catanduanes (cts)	✓	✓	✗
21. Bikol, Rinconada (bto)	✓	✓	✓
22. Bikol, Southern Catanduanes (bln)	✓	✓	✓
23. Binukid (bkd)	✓	✓	✓
24. Binukidnon	✓	✓	✓
25. Bisadya	✓	✓	✗

	Language	Sentence List	Word List	Audio Recording
26.	Bisaya	✓	✗	✓
27.	Blaan, Koronadal (bpr)	✗	✗	✓
28.	Blaan, Sarangani (bps)	✗	✗	✓
29.	Bogkalot (ilk)	✓	✓	✓
30.	Bolinao (smk)	✓	✓	✓
31.	Bontok, Central (lbk)	✓	✓	✓
32.	Bontok, Eastern (ebk)	✓	✓	✓
33.	Buhid (bku)	✗	✗	✓
34.	Butuanon (btw)	✓	✗	✗
35.	Capiznon (cps)	✓	✓	✗
36.	Cebuano (ceb)	✓	✓	✓
37.	Chavacano (cbk)	✓	✓	✓
38.	Cuyonon (cyo)	✓	✓	✓
39.	Davawenyo (daw)	✓	✓	✗
40.	Dumagat, Remontado (agv)	✓	✓	✓
41.	Ga'dang (gdg)	✓	✓	✓
42.	Gaddang (gad)	✓	✓	✓
43.	Hanunoo (hnn)	✗	✗	✓
44.	Hiligaynon (hil)	✓	✓	✓
45.	Ibaloi (ibl)	✓	✓	✓
46.	Ibanag (ibg)	✓	✓	✓
47.	Ifugao, Amganad (ifa)	✓	✓	✗
48.	Ifugao, Batad (ifb)	✓	✓	✗
49.	Ifugao, Mayoyao (ifu)	✓	✓	✓
50.	Ifugao, Tuwali (ifk)	✓	✓	✓
51.	Ifugao	✓	✓	✓
52.	Igorot	✓	✓	✓
53.	Ilocano (ilo)	✓	✓	✓
54.	Inabaknon (abx)	✗	✗	✓
55.	Inonhan (loc)	✓	✓	✓

Legacy Language Materials in the Ernesto Constantino Collection

Language	Sentence List	Word List	Audio Recording
56. Iranun (ilp)	✓	✓	✓
57. Iraya (iry)	✗	✗	✓
58. Isinay (inn)	✓	✗	✓
59. Isnag (isd)	✓	✓	✓
60. Itawit (itv)	✓	✓	✓
61. Itneg	✓	✓	✓
62. Itneg, Binongan (itb)	✓	✓	✓
63. Itneg, Inlaud (iti)	✓	✓	✓
64. Itneg, Maeng (itt)	✓	✓	✗
65. Itneg, Masadiit (tis)	✗	✗	✓
66. Itneg, Masadiit (tis)	✓	✓	✓
67. Itneg, Moyadan (ity)	✓	✓	✗
68. Ivatan (ivv)	✓	✓	✓
69. Kalagan (kqe)	✓	✓	✗
70. Kalagan, Kagan (kll)	✓	✓	✗
71. Kalanguya (kak)	✗	✗	✓
72. Kalinga	✓	✓	✓
73. Kalinga, Butbut (kyb)	✓	✓	✓
74. Kalinga, Limos (kmk)	✓	✗	✓
75. Kalinga, Lubuagan (knb)	✓	✓	✓
76. Kalinga, Tanudan (kml)	✓	✓	✓
77. Kalinga, Vanaw (bjx)	✓	✓	✓
78. Kamayo (kyk)	✓	✗	✓
79. Kankanaey (kne)	✓	✓	✓
80. Kapampangan (pam)	✓	✓	✓
81. Kinaray-a (krj)	✓	✓	✓
82. Maguindanaon (mdh)	✓	✓	✓
83. Mandaya (mry)	✓	✓	✓
84. Mangyan	✗	✗	✓
85. Manide (abd)	✓	✓	✓

	Language	Sentence List	Word List	Audio Recording
86.	Manobo	✓	✓	✓
87.	Manobo, Obo (obo)	✗	✗	✓
88.	Manobo, Cotabato (mta)	✗	✗	✓
89.	Manobo, Ata (atd)	✗	✗	✓
90.	Manobo, Dibabawon (mbd)	✓	✗	✓
91.	Manobo, Matigsalug (mbt)	✗	✗	✓
92.	Mansaka (msk)	✓	✗	✓
93.	Mapun (sjm)	✓	✓	✓
94.	Maranao (mrw)	✓	✓	✓
95.	Masbatenyo (msb)	✓	✓	✓
96.	Minamanwa (mmn)	✗	✗	✓
97.	Negrito	✗	✗	✓
98.	Palawano, Brooke's Point (plw)	✗	✓	✓
99.	Palawano, Central (plc)	✓	✓	✓
100.	Pangasinan (pag)	✓	✓	✓
101.	Parianon	✓	✓	✓
102.	Ratagnon (btn)	✓	✗	✓
103.	Romblomanon (rol)	✓	✓	✓
104.	Sama Balangingih (sse)	✓	✓	✓
105.	Sama, Central (sml)	✓	✓	✓
106.	Sama, Southern (ssb)	✓	✓	✗
107.	Samal	✓	✓	✓
108.	Sambal (xsb)	✓	✓	✓
109.	Sambal, Botolan (sbl)	✓	✗	✓
110.	Sambwangnon	✗	✗	✓
111.	Sangil (snl)	✗	✗	✓
112.	Sinama	✓	✓	✓
113.	Sorsoganon, Northern (bks)	✓	✓	✓
114.	Sorsoganon, Southern (srv)	✓	✓	✓
115.	Subanen, Northern (stb)	✓	✗	✓

Legacy Language Materials in the Ernesto Constantino Collection

Language	Sentence List	Word List	Audio Recording
116. Subanen, Southern (laa)	✗	✗	✓
117. Subanon	✓	✓	✓
118. Subanon, Kolibugan (skn)	✗	✗	✓
119. Sulod (srg)	✗	✗	✓
120. Surigaonon (sgd)	✓	✓	✓
121. Tagabawa (bgs)	✓	✓	✓
122. Tagakaulo (klg)	✓	✓	✓
123. Tagalog (tgl)	✓	✓	✗
124. Tagbanwa (tbw)	✗	✗	✓
125. Tagbanwa, Calamian (tbk)	✓	✓	✓
126. Tagbanwa, Central (tgt)	✓	✓	✗
127. Tandaganon (tgn)	✗	✗	✓
128. Tausug (tsg)	✓	✓	✓
129. Tawbuid (twb)	✗	✗	✓
130. Tboli (tbl)	✓	✓	✓
131. Teduray (tiy)	✓	✓	✓
132. Waray (war)	✓	✓	✓
133. Waray (war) or Baybayanon (bvy)	✓	✓	✗
134. Yakan (yka)	✓	✓	✓
135. Yogad (yog)	✓	✓	✗