# Hambaeon: Building a Comprehensive Akeanon Speech and Text Corpus Towards Language Preservation and Digital Inclusion

A Special Problem Proposal
Presented to
the Faculty of the Division of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Visayas
Miag-ao, Iloilo

In Partial Fulfillment
of the Requirements for the Degree of
Bachelor of Science in Computer Science by

VILLANUEVA, Joshua C.
VILLANUEVA, Mariefher Grace Z.

Francis D. DIMZON
Adviser

December 9, 2024

**Abstract**

Akeanon, a language spoken in Aklan, Philippines, is classified as a low-resource language (LRL) due to its limited linguistic resources and lack of digital integration. This research aims to develop and establish a comprehensive text and speech corpus, and build a model as a foundation for an automatic speech recognition (ASR) system for standardized Akeanon language. For the methodology, data collection will include compiling word lists for Akeanon based on the Swadesh 207 list and extracted text from an existing Akeanon-to-English dictionary. After compiling the word lists, their phonetic transcriptions will be encoded for the text corpus. For the development of the speech corpus, a selection of 1000 Akeanon words, having words from Swadesh list as priority, will be voice recorded from a total of fifty native speakers with varying gender and age. Audio resources containing the Akeanon language as well as their corresponding transcriptions that are available online will also be retrieved. The development of the speech and text corpus will be overseen and validated by a linguistics expert. A DNN model will be built and trained with Kaldi, using the newly developed corpus for initial results. This study aims to contribute to the preservation and digital inclusiveness of the Akeanon language and lay the groundwork for future works in developing an ASR system for the language.

**Keywords:** Language resources, Natural language processing (NLP), Speech recognition, Philippine languages, Aklan, Akeanon, Language corpus, Low-resource languages (LRL)

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Overview

Speech-to-Text (STT) technology has rapidly evolved in recent years, driven by advancements in deep learning algorithms such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), which have significantly improved the accuracy of STT systems (Televic, 2024). Open-source toolkits such as Kaldi have further accelerated research and development in this field by providing a flexible framework for building and training custom automatic speech recognition (ASR) models. ASR systems, which convert speech into text, have become essential components of various applications, from virtual assistants to transcription services (Cerna et al., 2023). However, despite these advancements, only a few Philippine languages have been explored and integrated into this technology. This special problem focuses on one of the understudied (Wellstood, 2022) Central Philippine languages, Akeanon.

Akeanon is an Austronesian language belonging to the Visayan subgroup (Biray, 2023). With more than 130,000 households (PSA, 2023) speaking the language, Akeanon is primarily spoken in the province of Aklan, located in northwestern Panay. Biray (2023) explains that the language has several dialects, each typically named after the town where it is spoken. These include Akeanon Buruangganon, Akeanon Nabasnon, Akeanon Bukidnon, and the standard Akeanon, which is spoken in Kalibo, the provincial capital of Aklan. For this special problem, the researchers will focus on developing the speech and text corpus specifically for standard Akeanon.

Up to this date, no studies have been conducted that is directly related to Akeanon and speech recognition altogether. However, there exist similar studies in the context of speech recognition on other regional languages such as Bisaya in the study of Cerna et al. (2023), Hiligaynon, studied by Billones and Dadios (2014), and in the study of Liao et al. (2019) for Bikol and Kapampangan. This special problem aims to bridge the gap in speech recognition for Akeanon starting with establishing a foundational speech corpus for the language, which can lay the groundwork for future research and applications. The corpus development will draw on methodologies from similar studies conducted for other regional languages such as the study of Cerna et al. (2023) and Liao et al. (2019), adapting them to meet the specific needs of Akeanon. In doing so, the project aims to bring Akeanon closer to digital integration, promoting inclusivity in speech recognition technology for Philippine languages. By bridging this gap, this special problem aspires to create a resource that can benefit future ASR developments, language preservation efforts, and the broader field of computational linguistics.

Creating a speech-to-text (STT) system for the Akeanon language not only fills

the gap in representation for this regional language but also aids in its preservation and fosters digital inclusion. This specific project aims to establish a foundational corpus that effectively captures the distinct speech patterns and intricacies of Akeanon, while taking into account the language's unique phonetic and linguistic features. Utilizing the resources gathered for this research, the team will concentrate on developing a comprehensive text and speech corpus that can provide a basis for future speech recognition systems pertaining to the Akeanon language. The researchers will also build and train a DNN model using Kaldi toolkit, to develop an ASR system that will provide initial speech recognition results for Akeanon. Finally, the study intends to investigate the challenges faced in developing speech models for languages with limited resources, offering valuable insights for the wider field of speech technology development.

## 1.2 Problem Statement

Akeanon remains underrepresented in modern speech technologies. According to Khan et al. (2023), in machine learning, natural language can be categorized into two categories: low-resource languages (LRLs) and high-resource languages (HRLs). Among these resources are (a) collections of text in different formats, such as research papers, journal articles, social media content, etc.; (b) lexical, syntactic, and semantic resources, such as dictionaries, bag of words, semantic databases, etc.; and (c) task-specific resources, such as annotated text, machine translation corpora, part-of-speech tags, etc.. HRLs e.g. English, French, Japanese, etc., are languages that are highly accessible and have many data resources that can be used for natural language processing (NLP). LRLs, on the

other hand, are understudied and have few data resources that can be utilized for NLP. Most regional languages in the Philippines are considered to be LRL, including the Akeanon language. Alejan et al. (2021) raised concerns on the Philippines' inclusion on a global list of the top ten "language hotspots", which means that many of its languages are disappearing faster than they are being completely documented. Their study noted the global rate of language extinction, which is one in every two weeks. They also projected that around half of the 6,000 languages will become extinct by the end of the century, to which most of them are indigenous languages. According to Magueresse et al. (2020) [29], a language supported by NLP techniques can help preserve it from extinction. It will also make the language more available and accessible in digital format, which offers significant commercial value, societal purpose, and applications in a variety of domains (Tsvetkov, 2017).

This special problem aims to address the lack of resources, availability, and accessibility of the Akeanon language in, but not limited to, modern speech technologies by building and establishing a speech and text and corpus for the language. Additionally, by developing an ASR model that is specific for Akeanon would lay the foundation for future research in speech-to-text, and other modern speech technologies for the language. Lastly, this special problem seeks to inspire innovation and drive similar efforts to preserve and develop accessible language technologies for other regional languages in the Philippines.

## 1.3 Research Objectives

### 1.3.1 General Objective

The general objective of this study is to construct and establish a comprehensive text and speech corpus for the Akeanon language, which can serve as a foundation for future development of language technologies and automatic speech recognition (ASR) systems. Additionally, the study aims to design and implement an ASR system for the language using the Kaldi toolkit.

### 1.3.2 Specific Objectives

Specifically, the study targets to:

1. Develop an Akeanon text corpus by collecting existing language resources such as dictionaries, word lists, and phonetic transcriptions, and organizing them into an annotated dataset.

2. Build a speech corpus by recording native speakers and using pre-existing Akeanon audio resources which can be found online.

3. Validate the text and speech corpus with the assistance of linguistic experts and native speakers to ensure accuracy and reliability.

4. Develop and evaluate an automatic speech recognition (ASR) model using deep neural networks (DNN) and the Kaldi toolkit with the newly created Akeanon corpora.

## 1.4 Scope and Limitations of the Research

The system is specific to the standard Akeanon language that is spoken in Kalibo, Aklan, the provincial capital. It is limited to the standard Akeanon even though the language has several dialects, due to resource constraints and the need for specificity. The study is centered around gathering audio samples from native speakers of standard Akeanon to guarantee precision and uniformity in the training of the model. This does not include other variations or dialects that are based on standard Akeanon language such as Buruanganon, Akeanon Nabasnon, and Akeanon Bukidnon, which can have different and unique phonetic and lexical traits. Nevertheless, the model's effectiveness might be influenced by the scarce availability of Akeanon data, potentially affecting its wide-ranging applicability.

## 1.5 Significance of the Research

Akeanon language, like many indigenous languages in the Philippines, lacks representation in digital technologies. Creating an automatic speech recognition (ASR) model for Akeanon language will help contribute to the preservation of the language in digital format, establishing a resource that will support documentation and education initiatives in the future. The dataset and model produced in the study of Akeanon language can act as a basis for further and additional linguistic research.

Akeanon and its incorporation in speech recognition technology fosters digital inclusivity. This enables Akeanon speakers to engage with technology in their

mother tongue highlighting the areas in education, communication, and public service where language barriers are almost present when accessing the said areas. Mobile applications, virtual assistants and other tools can embed speech to text of Akeanon language to help enhance accessibility and boost engagement with technology.

The challenge faced and lessons learned from this study will help contribute to addressing the lack of representation of low-resource language in AI technology, aligning with the need for inclusivity in language processing (Poupard, 2024). This initiative will help in promoting linguistic diversity as well as safeguard cultural heritage through Akeanon speech recognition in technological advancement. Poupard (2024) highlights that even minimal focus on languages with fewer resources can significantly influence their viability in an increasingly digital world where larger languages prevail.

# Chapter 2

# Review of Related Literature

## 2.1 Automatic Speech Recognition

Automatic Speech Recognition (ASR) is a technology that processes human speech into readable text by the use of machine learning or artificial intelligence (AI). The ASR system has grown popular over the past decade as it quickly approaches human accuracy levels, there is a great demand for applications taking advantage of ASR technology in their products to make audio and video data more accessible (Foster, 2023).

Automatic Speech Recognition independently decodes and transcribes spoken language using a machine-base process. An ASR system takes in acoustic signals from a speaker via a microphone, analyzes these signals using various patterns, models, or algorithms, and generates an output, most commonly in text form (Levis & Suvorov, 2012). The importance of differentiating speech recognition

from speech understanding (speech identification) is that, speech understanding focuses on interpreting the meaning of an utterance rather than merely transcribing it. Furthermore, speech recognition is distinct from voice recognition: speech recognition pertains to a machine's capability to identify the words spoken, while voice recognition relates to a machine's ability to discern the manner of speaking (Levis & Suvorov, 2012).

## 2.2 Lexicon Model

The lexicon model is essential in automatic speech recognition, serving as the bridge between the acoustic representation and the sequence of words produced by the speech recognizer. The lexicon's function can be viewed in two aspects: it first identifies the words or lexical items recognized by the system, and second, it offers the framework to develop acoustic models for each entry (Adda-Decker & Lamel, 2000). Consequently, lexical design consists of two primary components: determining and selecting the vocabulary items and representing each pronunciation entry using the fundamental acoustic units of the recognizer. In large vocabulary speech recognition, the vocabulary is typically chosen to optimize lexical coverage within a specified size of the lexicon, and the basic units selected are generally phonemes or phone-like units ((Adda-Decker & Lamel, 2000).

## 2.3 Acoustic Model

Acoustic modeling is a fundamental and preliminary step in the process of speech recognition. The acoustic model defines the relationship between acoustic data and linguistic elements. Most calculations in acoustic modeling are attributed to feature extraction and statistical representation, making it a crucial factor in the recognition process. Statistical representations are derived from the features that have been extracted (Bhatt et al., 2020). In the acoustic model, the distribution of these extracted features corresponding to specific sounds is modeled to create a connection between the features and the structures of the linguistic units.

According to Bhatt et al. (2020), several techniques for feature extraction, including those based on human perception and the mechanics of voice production, have been documented. Features were derived for acoustic modeling in a speaker-independent recognition context since such systems pose challenges in speech recognition.

## 2.4 Language Model

Language models are crucial for various daily applications, including correcting grammatical errors, recognizing speech, and summarizing text. Due to the recent advancements in deep learning techniques, conventional n-gram and word embedding language models are being substituted with neural network-based models (Mago & Qudar, 2020).

Large Language Models (LLMs) have recently shown remarkable abilities, en-

compassing tasks like natural language processing (NLP), language translation, text generation, and answering questions. In addition, LLMs play a vital role in computerized language processing, capable of grasping intricate verbal patterns and producing relevant and coherent responses in various contexts. However, the significant advancements in LLMs have led to a surge in research contributions, making it challenging to fully comprehend the overall impact of these developments (Fahad et al., 2024).

## 2.5 Local Dialects and Low-Resource Languages On Automatic Speech Recognition

Deep learning technologies have evolved from rudimentary systems to advanced models that can fluently comprehend natural language, making remarkable progress in their integration into Automatic Speech Recognition (ASR). Neural networks have become crucial in ASR for capturing temporal dynamics and phonetic differences, enabling wider use in virtual assistants, educational applications, and customer support (Alharbi et al., 2021). Noisy environments where background sounds significantly impair the accuracy and dependability of speech recognition. The considerable challenge for languages with limited resources is the size of the vocabulary. This influences the performance of the model in which larger vocabularies enhance adaptability but demand more data and computational power. ASR systems struggle with dialectal variation, which can impede model accuracy due to differences in pronunciation, a concern for languages such as Akeanon, known for its various dialects (Alharbi et al., 2021).

## 2.6 The Akeanon Language

### 2.6.1 History and its Speakers

Zorc (1995) stated that Akeanon serves as the main language in the northwestern area of Panay Island in the central Philippines, boasting over 350,000 speakers. Both the language and its speakers derive their name from the Akean River, which runs through the heart of the province by the same name. The people, culture, and items linked to this river and region are referred to as Aklanon, while the language is known as Inakeanon, incorporating the -in- infix and an accent alteration, or more generally Bisaya, as Aklanons identify themselves as part of the Visayan cultural and linguistic family. Many Aklanons, particularly those in professional fields, have relocated to various major cities in the Philippines, such as Manila, Cebu, and Cotabato, in pursuit of job opportunities, with sizable communities also found in San Francisco and New York. The dialect discussed here is that of Kalibo, Aklan, the provincial capital and its main commercial hub. Other dialects are linked to the towns of Altavas, Batan, Balete, Banga, Madalag, New Washington, Numancia, Malinao, Lezo, Makato, Tangalan, Nabas, Ibajay, and Libacao—though the latter two show significant divergence, they remain mutually understandable with the others. Two towns exist within Aklan province that feature different dialects—with Buruanga associated with Kinaray-a, and Malay linked to various dialects of Tablas, Romblon. The closest languages to Akeanon are Kinaray-a and Kuyonon, both of which belong to the West Bisayan subgroup of Central Philippine languages.

### 2.6.2 Phonology

#### Akeanon Phonology: Historical and Synchronic Perspectives

The Akeanon language, native to the Aklan province in the Philippines, possesses a distinctive phoneme that sets it apart from other Philippine-type languages. Initially recognized as a voiced velar fricative and subsequently categorized as a velar approximant, this phoneme differentiates Akeanon from its linguistic siblings within the Bisayan group, such as Hiligaynon, Cebuano, and Kinaray-a. The initial examinations of this phoneme can be traced back to Scheerer (1920), who analyzed its correlation with Bisayan cognates. Subsequent research by de la Cruz and Zorc (1968) characterized it as a voiced velar fricative, functioning both as a consonant and a semivowel. More recent studies have reiterated its classification as a velar approximant, emphasizing its absence of articulatory turbulence (Zorc, 2005; Rentillo & Pototanon, 2022)

#### Linguistic Status and Usage of Akeanon

Akeanon is acknowledged as an institutional language according to the Expanded Graded Intergenerational Disruption Scale (EGIDS) and is included in the Mother Tongue-Based Multilingual Education (MTB-MLE) program in primary education. With approximately 500,000 speakers based on recent estimates, the language flourishes in both spoken and written forms, encompassing social media, radio programs, and public signages. Its phonological framework, which is defined by a three-vowel inventory and distinctive consonantal reflexes, has been influenced by historical changes and cross-linguistic interactions.

#### Crosslinguistic Comparisons and Historical Accounts

The evolution of the Akeanon phoneme is believed to reflect more extensive linguistic trends, such as velarization and palatalization, seen in various languages. Rentillo and Pototanon (2022) contend that the development of the phoneme may have been shaped by regional linguistic changes or historical interactions with other Bisayan dialects. Moreover, historical accounts from figures such as de Méntrida (1841) and Monteclaro (1957) indicate cultural and linguistic connections to Borneo, which influenced the distinct characteristics of Akeanon speech.

### Acoustic and Articulatory Characteristics

Recent acoustic studies conducted by Rentillo and Pototanon (2022) offer empirical insights that differentiate the velar approximant from other phonemes. Their research demonstrates that the formant frequencies (F1 and F2) of this phoneme are lower than those of vowels, with variations that depend on adjacent phonological contexts. These findings emphasize the phoneme's unique articulatory properties, confirming its classification as an approximant rather than a fricative.

### Implications for Language Documentation

The distinctive attributes of Akeanon phonology reinforce the significance of documenting endangered and lesser-known languages. The Akeanon phoneme acts as a case study for exploring phonological diversity and innovation within Philippine languages. As noted by Rentillo and Pototanon (2022), further research could yield greater understanding of the historical and sociolinguistic elements that influence such unique linguistic features.

### 2.6.3 Morphology

#### Morphology and its Role in Language

Morphology, which examines word structures and their smallest meaningful units, is fundamental to comprehending the formation and development of languages. In various languages, including Akeanon, derivational morphology transforms syntactic roles or introduces novel meanings through methods like affixation, reduplication, subtraction, and internal modification of words. These methods not only redefine lexical meanings but also influence word categories like parts of speech (Biray, 2023).

#### Linguistic Diversity in the Philippines

The Philippines is distinguished by its extensive linguistic variety, containing over 180 distinct languages, predominantly of Austronesian origin. Akeanon, which has approximately 460,000 speakers, belongs to the Malayo-Polynesian language family and functions as an official language in the province of Aklan. The language shares lexical similarities with Kinaray-a and Kuyunon, accompanied by notable dialectical variations throughout the area.

#### Akeanon Dialectical Variations

Akeanon dialects—including Standard Akeanon, Buruangganon, Nabasnon, and Bukidnon—display specific linguistic characteristics. These dialects are shaped by their geographical and cultural backgrounds, resulting in differences in structure, word order, and affixation. For example, reduplication serves as a prominent morphological feature that modifies meanings, whereas circumfixes are frequently

utilized for the formation of new words. Dialect-specific phonemic variations, such as replacing "l" with "r" in certain instances, further highlight these distinctions.

**Social and Cultural Significance**

The Akeanon language mirrors the social traits of its speakers, showcasing values such as hospitality and respect. Expressions of endearment and polite language are prevalent in daily interactions, emphasizing the cultural identity of the community. Despite structural differences, the fundamental meanings of expressions remain uniform across dialects, illustrating the language's strength and flexibility.

**Challenges and Preservation Efforts**

Like many other languages in the Philippines, Akeanon faces challenges stemming from modernization and the growing impact of technology. Initiatives to safeguard the language include its integration into the Mother Tongue-Based Multilingual Education (MTB-MLE) framework and the creation of orthographies that document its linguistic characteristics. Nonetheless, further support from both local and national organizations is crucial to maintain and promote the language in the face of the rising influence of global languages.

### 2.6.4 The 300 Languages Project: A Worldwide Linguistic Initiative

The 300 Languages Project, led by The Rosetta Project and The Long Now Foundation, stands as a groundbreaking effort aimed at creating a universal collection of human languages. This project seeks to gather and digitize parallel text and

audio data from the 300 most frequently spoken languages around the globe. This extensive initiative addresses the significant shortage of resources for linguistic research, particularly for lesser-known languages, by utilizing volunteer-submitted public domain texts and recordings, all of which will be made available through The Internet Archive.

### Linguistic Variety and Digital Visibility

Among the roughly 7,000 languages spoken worldwide, merely 20-30 languages possess a substantial digital footprint, including English, Spanish, and Mandarin. These languages, in conjunction with the next 270-280 most spoken languages, encompass over 90% of the global populace. In contrast, the remaining 10% communicate in one of the 6,700 minority languages, many of which are at risk of extinction due to inadequate digital and physical documentation. The 300 Languages Project highlights the importance of showcasing these minority languages by establishing a scalable "seed corpus" that begins small but is intended to expand sustainably.

### Contributions to Multilingual Research and Technological Advancements

This initiative distinguishes itself by merging linguistic preservation with technological innovation. By assembling a large-scale public domain multilingual parallel corpus, the project enables progress in speech recognition, automated translation, and cross-linguistic studies. The absence of such resources has historically limited research and development to a small number of languages with existing corpora. The project's focus on widely translated texts, such as the Swadesh List, the Universal Declaration of Human Rights, and chapters 1-3 of Genesis, ensures extensive

applicability for linguistic research and tech applications.

### Volunteer-Driven, Scalable Approach

The project's dependence on volunteer-contributed materials highlights its scalability and cost-efficiency. By establishing a comprehensive protocol for language documentation, this effort lays out a replicable model for documenting additional languages beyond the initial 300. The low-cost, community-focused method reflects earlier successful documentation endeavors like the ancient Rosetta Stone, which facilitated the understanding of Egyptian hieroglyphs through parallel texts.

### Significance for Language Conservation

The 300 Languages Project plays a crucial role in preserving linguistic diversity by documenting and archiving minority languages that are on the brink of disappearing. By making multilingual resources publicly accessible, the initiative not only benefits researchers but also bolsters educational and cultural preservation efforts worldwide. Its alignment with the ALLOW initiative at the Language Technologies Institute further demonstrates a collaborative dedication to advancements in speech and language technologies.

# Chapter 3

# Research Methodology

This chapter discusses the methodology used to develop the text and speech corpus for the Akeanon language, as well as building, training, and testing a model to generate initial results. The chapter will be divided into two main parts: Research activities and the calendar of activities for this special problem.

## 3.1 RESEARCH ACTIVITIES

Figure 3.1 shows the general overview of the methodology for the development of an ASR system for the Akeanon language.

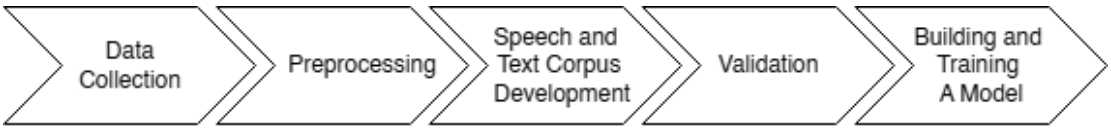


Figure 3.1: Research Methodology

### 3.1.1 Data Collection

#### Collating Pre-existing Online Resources

For the data collection, the researchers will make use of existing online resources from the website, Bible.com (see Appendix A for permission to use their resources). These resources include recordings and transcriptions of the Akeanon translations of the multiple books and chapters of the Bible.

#### Compiling Akeanon Words

The researchers will collect equivalent Akeanon words based off on the Swadesh 207 word-list and use the Aklanon to English Dictionary by Zorc, Reyes, and Prado (1969) as reference. In addition to the Swadesh list, the researchers have reached out to Dr. David Zorc, one of the authors of the Aklanon-to-English Dictionary. Not only who he granted the researchers permission to use the dictionary, but he has also generously provided a soft copy of an existing database containing all the contents of the dictionary to the researchers. This database was shared courtesy of Dr. Jarrette Allen, a linguist who has also worked on the Akeanon language.

#### Phonetic Transcription

After compiling the Akeanon word lists, their phonetic transcription will also be encoded, using the work of Rentillo and Pototanon (2022) as reference for Akeanon phonology. Audio transcriptions retrieved from the Bible.com will also have their phonetic transcriptions be encoded. For the phonetic transcription the researchers will seek assistance from Ms. Hazel Cipriano, a linguist who is also a native speaker of the language.

### Word Selection for Speech Corpus

For building the speech corpus, the researchers will prioritize words from the Swadesh 207 list for the voice recordings. Additionally, to complete the selection of 1000 words that will be prepared for speakers to read, the researchers will collaborate with a linguistic expert during the selection.

### Voice Recording

A total of 50 native speakers of standard Akeanon will be gathered for the recording of the generated word list. The researchers have also sought a collaboration with Aklan State University (ASU) - College of Teacher Education for the selection of speakers. The speakers will be of varying gender and age. Table 3.1 shows the categories of native speakers. For the audio recording, the microphone that will be used will be Shure SM58 (dynamic, cardiod pick-up pattern) with a Focusrite Scarlett 2i2 audio interface, having Adobe Audition 2021 as the recording software. In case of unavailability, recording over a smartphone with noise-cancelling headphones would suffice. The audio file will be named in the following convention: <gender>_<age_group>_<speaker_number>.

Table 3.1: Categories of Native Speakers

| Category | Subcategories |
|---|---|
| Gender | Male<br>Female<br>Non-binary<br>Others |
| Age Group | 12-15<br>16-30<br>31-45<br>46-65 |

### 3.1.2 Preprocessing

For preprocessing audio files, Adobe Audition 2021 will be used for recording and audio processing, which will include normalization and noise reduction of the recorded audio. The cleaned up audio files will then be exported in a WAV format.

### 3.1.3 Speech and Text Corpus Development

The previous steps will set as a precedent for the speech and text corpus development. The development of the text corpus for the Akeanon language involves creating a comprehensive collection of audio transcriptions from pre-existing online resources, dictionary-based word list, and their respective phonetic transcriptions. For the speech corpus, the collated audio transcriptions and word list will be mapped with their corresponding voice recordings and will be annotated accordingly.

### 3.1.4 Validation

In collaboration with ASU College of Teacher Education and another linguistics expert, they will be aiding the researchers by validating the newly created text and speech corpora.

### 3.1.5 Building and Training A Model

To generate initial results for an ASR system, the researchers will build, train, and test a model using the Kaldi toolkit. Similar to the methods employed by Panizales et al. (2023) A six-fold cross validation scheme will be used for testing and training the data. Finally, the researchers have decided to use a DNN model for training.

## 3.2 Calendar of Activities

Table 3.2 shows a Gantt chart of the activities. Each bullet represents approximately one week worth of activity.

Table 3.2: Timetable of Activities

| Activities (A.Y. 2024-2025) | Sep | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May |
|---|---|---|---|---|---|---|---|---|---|
| Brainstorming and Selection of Topic | ● | ●● | | | | | | | |
| Drafting and Finalization of Chapter 1 - Introduction | | ●● | ● | | | | | | |
| Drafting and Finalization of Chapter 2 - Review of Related Literature | | | ●●● | | | | | | |
| Preparation of Letters; UP-VREB and Establish Communication for Collaboration | | ● | | | | | | | |
| Drafting and Finalization of Chapter 3 - Methodologies | | | ● | ●● | | | | | |
| Proposal Document Creation in LaTex | | | ● | ● | | | | | |
| Proposal Presentation | | | | ● | | | | | |
| Data Gathering | | | | ●●● | ●●● | | | | |
| Preprocessing | | | | | ● | ●●● | | | |
| Drafting and Finalization of Chapter 4 - Results and Discussion | | | | | | ● | ●● | | |
| Corpus Development and Validation | | | | | | | ●●● | ●●● | |
| Drafting and Finalization of Chapter 5 - Summary and Recommendation | | | | | | | | ●● | ●● |
| Drafting and Finalization of SP Defense | | | | | | | | | ● |
| SP Defense | | | | | | | | | ● |

# References


Adda-Decker, M., & Lamel, L. (2000). The use of lexica in automatic speech recognition. In F. Van Eynde & D. Gibbon (Eds.), *Lexicon development for speech and language processing* (pp. 235–266). Dordrecht: Springer Netherlands. Retrieved from https://doi.org/10.1007/978-94-010-9458-0_8 doi: 10.1007/978-94-010-9458-0_8

Alejan, J. A., Ayop, J. I. E., Allojado, J. B., Abatayo, D. P. B., Abacahin, S. K. N., & Bonifacio, R. (2021, May). *Heritage language maintenance and revitalization: Evaluating the language endangerment among the indigenous languages in bukidnon, philippines.* Retrieved from https://eric.ed.gov/?id=ED617996 (ERIC - Online Submission)

Alharbi, S., Alrazgan, M., Alrashed, A., AlNomasi, T., Almojel, R., Alharbi, R., ... Almojil, M. (2021, 09). Automatic speech recognition: Systematic literature review. *IEEE Access*, *PP*, 1-1. doi: 10.1109/ACCESS.2021.3112535

Bhatt, S., Jain, A., & Dev, A. (2020, 01). Acoustic modeling in speech recognition: A systematic review. *International Journal of Advanced Computer Science and Applications*, *11*. doi: 10.14569/IJACSA.2020.0110455

Billones, R. K. C., & Dadios, E. P. (2014). Hiligaynon language 5-word vocabulary speech recognition using mel frequency cepstrum coefficients and genetic al-

gorithm. In *2014 international conference on humanoid, nanotechnology, information technology, communication and control, environment and management (hnicem)* (p. 1-6). doi: 10.1109/HNICEM.2014.7016247

Biray, E. (2023, 12). Derivational morphology features in common akeanon dialects. *International Journal of Language and Literary Studies*, *5*, 222-234. doi: 10.36892/ijlls.v5i4.1441

Cerna, P. D., Cascaro, R. J., Juan, K. O. S., Montes, B. J. C., & Caballero, A. O. (2023). Bisayan dialect short-time fourier transform audio recognition system using convolutional and recurrent neural network. *International Journal of Advanced Computer Science and Applications*, *14*(3). Retrieved from `http://dx.doi.org/10.14569/IJACSA.2023.01403111` doi: 10.14569/IJACSA.2023.01403111

Fahad, N. M., Fatema, K., Mukta, S., & Raiaan, M. A. K. (2024). A review on large language models: Architectures, applications, taxonomies, open issues and challenges. *Computer Science*. Retrieved from `https://www.mdpi.com/2227-7390/11/21/4493` doi: 10.1109/ACCESS.2024.3365742

Foster, K. (2023). *What is automatic speech recognition? a comprehensive overview of asr technology.* Retrieved from `https://www.assemblyai.com/blog/what-is-asr/` (AssemblyAI)

Khan, M., Ullah, K., Alharbi, Y., Alferaidi, A., Alharbi, T. S., Yadav, K., ... Ahmad, A. (2023). Understanding the research challenges in low-resource language and linking bilingual news articles in multilingual news archive. *Applied Sciences*, *13*(15). Retrieved from `https://www.mdpi.com/2076-3417/13/15/8566` doi: 10.3390/app13158566

Levis, J., & Suvorov, R. (2012, 11). Automatic speech recognition.. doi: 10.1002/9781405198431.wbeal0066

Liao, E., Ganareal, K., Paguia, C., Agreda, C., Octaviano, M., & Rodriguez, R. (2019, 11). Towards the development of automatic speech recognition for bikol and kapampangan. In (p. 1-5). doi: 10.1109/HNICEM48295.2019.9072783

Mago, V., & Qudar, M. (2020). *A survey on language models.* Retrieved from https://www.researchgate.net/publication/344158120_A_Survey_on_Language_Models3

Magueresse, A., Carles, V., & Heetderks, E. (2020). Low-resource languages: A review of past work and future challenges. *CoRR*, *abs/2006.07264*. Retrieved from https://arxiv.org/abs/2006.07264

Panizales, J. P., Jr., B. G., & Piorque, L. (2023). *Speaknow: A speech-to-text system for the hiligaynon language using kaldi toolkit.* Undergraduate Thesis, University of the Philippines Visayas. (Accessible through the UPV Computer Science Faculty)

Poupard, D. (2024). Attention is all low-resource languages need. *Translation Studies*, *17*(2), 424–427. Retrieved from https://doi.org/10.1080/14781700.2024.2336000 doi: 10.1080/14781700.2024.2336000

PSA. (2023). *Tagalog is the most widely spoken language at home (2020 census of population and housing).* Retrieved from https://psa.gov.ph/content/tagalog-most-widely-spoken-language-home-2020-census-population-and-housing

Rentillo, P., & Pototanon, R. M. D. (2022, Jan.). A synchronic and historical look at aklanon phonology. *Acta Linguistica Asiatica*, *12*(1), 91–127. Retrieved from https://journals.uni-lj.si/ala/article/view/10359 doi: 10.4312/ala.12.1.91-127

Televic. (2024, 1). *The evolution of speech-to-text technology.* Re-

trieved from https://www.televic.com/en/televicgsp/news/the-evolution-of-speechtotext-technology

Tsvetkov, Y. (2017). *Opportunities and challenges in working with low-resource languages.* Retrieved from https://www.cs.cmu.edu/~ytsvetko/jsalt-part1.pdf (PDF)

Wellstood, Z. (2022). A relative clause analysis of event existential constructions in aklanon. *GLOSSA*, *7*(1). Retrieved from https://www.glossa-journal.org/article/id/5866/ doi: 10.16995/glossa.5866

Zorc, R. D. (1995). Aklanon r. david zorc. In D. T. Tryon (Ed.), *Comparative austronesian dictionary* (pp. 343–350). Berlin, New York: De Gruyter Mouton. Retrieved from https://doi.org/10.1515/9783110884012.1.343 doi: doi:10.1515/9783110884012.1.343

Zorc, R. D., Reyes, V. S., & Prado, N. (1969). A study of the aklanon dialect, volume two: Dictionary (of root words and derivations), aklanon to english..