UNIVERSITY OF THE PHILIPPINES VISAYAS
COLLEGE OF ARTS AND SCIENCES
DIVISION OF PHYSICAL SCIENCES AND MATHEMATICS

CMSC 198.1 Special Problem I
First Semester AY 2022-2023

Filipino Automatic Speech Recognition for Filipino Children Aging six to 14

Chapter 2
Review of Related Literature

Submitted by
Michael Vincent Dosado, 2019-04469;
Kevin Christian Lao, 2019-12279;
Allen Solomon Tam, 2019-03019
B.S. in Computer Science IV
College of Arts and Sciences
University of The Philippines Visayas
Miagao, Iloilo

Submitted to
Almie P. Carajay
Asst. Prof., Faculty of Computer Science
Division of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Visayas
Miagao, Iloilo

CHAPTER 2

REVIEW OF RELATED LITERATURE

This chapter provides an overview of related literature, studies, and works that share the same theoretical background, research questions, proposed solutions, and/or end results as this publication. In order to highlight the similarities and differences between the linked studies and this work, the researchers reviewed extensive amounts of literature and studies from both national and international sources online.

Automated Speech Recognition

Speech recognition or speech to text, is a technology that recognizes speech data, allowing the voice as the main means of communication with the computer. Speech recognition apps have been a huge help for writers to boost their writing production and give opportunities for aspiring writers with physical impairements. Speech Recognition combines the fields of linguistics, math, and statistics. It detects human speech and translates it into text, and its aspects include a customized dictionary, acoustic modeling, pronunciation modeling, a decoder, and output - these are to be considered in creating Speech recognition technology (IBM Cloud Education, 2020).

Furthermore, speech recognition applications are done in three process: acoustic processing, feature extraction, and output recognition. In acoustic processing, speech inputs need to be processed to remove irrelevant noises in the background. Then, the task of feature extraction is to form computational representation of these speech waves as processing information in acoustic signal can be difficult to deal with. Extraction of efficient features derived desirable recognition of speech. Therefore, the goal of feature

extraction is to derive reliable and logical representations from cumbersome speech signals (Alim & Rashid, 2018).

Approaches to Speech Recognition

Acoustic Phonetic Approach

Before the emergence of the feature algorithms widely used today, the premature phase of speech recognition depended mainly on recognizing speech sound and find appropriate labels to them. This approach is based on a hypothesis, which states that there exist finite phonetic units that are uniquely different from each other, and these units are identified by alternating acoustic signals in regards to its time. In addition, these message carriers of speech are influenced by binary acoustic properties such as nasality, fricative sound,background noise, and continuous change in frequencies (Emeeshat, 2017).

Pattern Recognition Approach

Through use of well designed framework, this approach trains the machine to identify speech pattern representation in given set of speech signals as training models. Then, it compares each new speech to the speech pattern representations created in training stage to identify the unknown speech signal by using the matching algorithm (Emeeshat, 2017).

Artificial Intelligence Approach

This approach is a combination of both acoustic phonetic approach and the pattern recognition approach. The acoustic phonetic and pattern recognition lacks the capability to examine noticeable insight into human speech processing. The algorithm is assisted by knowledge in order to enhance the performance, and plays a key role in choosing an input representation that is suitable, the definition of units of speech, or the algorithm's structure itself. This method was found to be effective in solving the challenge brought by inter and intra speaker speech variability, and the ability to create decision trees (Emeeshat, 2017).

Developing a Children's Filipino Speech Corpus (CFSC) for Application in Automatic Detection of Reading Miscues and Disfluencies

In this paper, the researchers designed a Speech Corpus (speech database) on the Filipino Speech of Filipino Children. They have also presented the CFSC design, reading text, data collection procedure and speech transcription method. This was due to Speech Corpus having already been developed for languages such as English, Dutch, Chinese Mandarin, Italian, German and Swedish. In the case of Filipino, having features and orthography, there have been no speech corpus developed yet. The researchers aimed for the development of computer-assisted oral reading assessment and learning systems such as reading miscue detector (RMD) and automated reading tutor (ART) which can be assisted by the developed Speech Corpus (CFSC) to increase the quality of learning given by these systems.

Automatic Speech Recognition (ASR) Systems for Children: A Systematic Literature Review

In this paper, the researchers designed an Automatic Speech Recognition (ASR) on the Speech of Children. This was due to a realization wherein ASRs had already been developed for adult age. In the case of children's language, there have been no significant developments yet. The researchers aimed to throw light on the trends of research in children's speech recognition and analyze the potential of trending techniques to recognize children's speech. They have stated that it is challenging to gather a corpus of children's speech for the ASR system, and the data sets are typically smaller than the adult corpus. Adults and children also have diverse linguistic and acoustic characteristics, such as varied spectral and temporal features. There is a mismatch between children's and adult speech as a result of these differences. In order to offer a systematic literature review (SLR), the pertinent data from 76 studies on ASR for kids that were published between 2009 and 2020 were extracted. This review's goal is to shed light on the most recent developments in the field of children's speech recognition research and to assess the possibilities of emerging methods for doing so.

Assistive Mobile App for Children with Hearing & Speech Impairment Using Character and Speech Recognition

In this paper, the researchers designed an Assistive Mobile App that uses Character and Speech Recognition for Children with Hearing & Speech Impairment . This was as a result of the researchers' data collection, which revealed that the Philippines' public school system is still unable to serve 97 percent of children with disabilities (CWDs). This means that even if there is a national law requiring inclusion of

CWDs and the right to a basic education, only a small number of students have access to special education.

The researchers used the technological advancement of speech and character recognition in general education to provide a supplemental tool for preschool students with speech and hearing impairment under Special Education (SPED) class in learning basic counting, English alphabet, and recognizing basic shapes. The created system was put to the test by SPED instructors, parents, and the intended audience, and it has proven to be a useful tool for improving kids' writing and speaking skills.

Web-Based automation speech-to-text application using audio recording for meeting speech

In order to make the recording of meeting materials more effective and efficient, the researchers in this study construct a web-based automation speed-to-text tool that can record meeting participants' voices and convert them into text automatically. This system can be used by utilizing the voice recognition feature, known as Web Kit Speech Recognition. The average value for the duration system for the languages of Indonesia and English is 96,63% and 82,78% after successfully implementing this application, which is known as the Speech Meeting Web kit system.

Speech Recognition for Educational Entertainment

Globally, youth internet users has increased by 40% in the prior pandemic 2019. Today, experts have suggested that children in the primary age have a relative screen time of that of an adult - spending about more than five hours per day on the screens (Scanlon, 2020). Simultaneously, Filipino children at the primary have trouble in reading. Nine out of 10 Filipino students have low proficiency in reading - and this has been

looming the basic education system of our country (Oseña-Paez, 2022). Additionally, the Philippines placed the lowest in the ranking when it comes to reading comprehension of students, according to the 2018 Programme for Internation Student Assessment(San Juan, 2019).

The team found it timely and productive to help in addressing the learning crisis on our basic education system, specifically the lack of Filipino reading and writing skills, thus the team aims to develop an English-Filipino Automated Speech Recognition for Filipino children in the primary age.

References

Alim, S. A., & Rashid, N. K. A. (2018, July 20). *Some Commonly Used Speech Feature Extraction Algorithms*. InTechOpen. Retrieved November 6, 2022, from https://www.intechopen.com/chapters/63970

Bhardwaj, V., Ben Othman, M. T., Kukreja, V., Belkhier, Y., Bajaj, M., Goud, B. S., Rehman, A. U., Shafiq, M., & Hamam, H. (2022, April 27). *Automatic speech recognition (ASR) systems for children: A systematic literature review*. MDPI. Retrieved November 8, 2022, from https://www.mdpi.com/2076-3417/12/9/4419/htm

Emeeshat, J. S. (2017, August). Isolated Word Speech Recognition System For Children With Down Syndrome. 4-6.

IBM Cloud Education. (2020, September). *Speech Recognition*. IBM. Retrieved September 21, 2022, from https://www.ibm.com/cloud/learn/speech-recognition#:~:text=Speech%20recognizers%20are%20made%20up,to%20determine%20the%20appropriate%20output

Mary Jane C. Samonte Mapua University, Samonte, M. J. C., University, M., Renz Jirhel D. Bahia Mapúa University, Bahia, R. J. D., University, M., Samantha Bernadine A. Forlaje Mapua University, Forlaje, S. B. A., John Gabriel J. Del Monte Mapua University, Monte, J. G. J. D., Jecelle Anne J. Gonzales Mapua University, Gonzales, J. A. J., University, M. V. S. M., Sultan, M. V., & Metrics, O. M. V. A. (2018, October 1). *Assistive mobile app for children with Hearing & Speech Impairment using character and speech recognition: Proceedings of the 4th International Conference on Industrial and Business Engineering*. ACM Other conferences. Retrieved November 8, 2022, from https://dl.acm.org/doi/abs/10.1145/3288155.3288182

Oseña-Paez, D. (2022, August 10). *Why 9 out of 10 Filipino children can't read*. The Manila Times. Retrieved October 11, 2022, from

https://www.manilatimes.net/2022/08/10/opinion/columns/why-9-out-of-10-filipino-children-cant-read/1854046

Ronald Pascual & Rowena Cristina L. Guevara. (2012, November). *Developing a children's Filipino speech corpus for application in ...* https://www.researchgate.net/publication/261025285_Developing_a_children's_Filipino_speech_corpus_for_application_in_automatic_detection_of_reading_miscues_and_disfluencies. Retrieved November 3, 2022, from https://www.researchgate.net/publication/261025285_Developing_a_children's_Filipino_speech_corpus_for_application_in_automatic_detection_of_reading_miscues_and_disfluencies

Samonte, M. J. C., University, M., University, N. T., University, T. C., University, N. C. K., & Metrics, O. M. V. A. (2019, February 1). *Speech recognition and Filipino sign language E-tutor system: Proceedings of the 2nd International Conference on Image and Graphics Processing*. ACM Other conferences. Retrieved November 3, 2022, from https://dl.acm.org/doi/abs/10.1145/3313950.3313970

San Juan, R. (2019, December 3). *Philippines lowest in reading comprehension among 79 countries*. Philstar. Retrieved October 8, 2022, from

https://www.philstar.com/headlines/2019/12/03/1974002/philippines-lowest-reading-comprehension-among-79-countries

Yasir1, M., Nababan1, M. N. K., Laia1, Y., Purba1, W., Robin2, & Gea3, A. (2019, July 1). *IOPscience*. Journal of Physics: Conference Series. Retrieved November 8, 2022, from https://iopscience.iop.org/article/10.1088/1742-6596/1230/1/012081