

Reading Fluency Evaluation for Malaysian Primary School Children Using Feature Extraction Techniques in Speech Recognition

Abu Bakar, N.A.¹, Ahmad, R.², Sarlan, A.³

^{1,2,3} Computer and Information Science Department, University of Technology PETRONAS, Perak, Malaysia

²rohiza_ahmad@utp.edu.my

ABSTRACT

It is no longer compulsory for lower primary school children to sit for examinations, they will now be evaluated through a more objective assessment, known as PBD (classroom-based assessment). Though it may seem to benefit both teachers and students in reducing stresses and anxiety over sitting for exams, but it would cause teachers to be burdened by the extremely high work-loads due to the frequent assessment. This could result in extreme pressure for the teachers which would eventually reach to the point of non-objectivity as well as inconsistency in students' evaluation. Hence, this paper intends to investigate on the ways of easing the teachers by exploring the subject of speech recognition technique with the ability to evaluate students' fluency based on a standardized set of criteria in reading assessment. Initial study conducted shows that the technique that is frequently used in speech recognition is Mel-Frequency Cestrum Coefficient (MFCC) method which was found to be able to extract the features needed in evaluating the fluency of a speech. Moreover, these extracted features are able to match with the criteria of a fluent reading which are pausing, phrasing, intonations, accuracy as well as pace of reading.

Keywords

Feature extraction techniques, reading fluency evaluation, speech recognition, standard set of fluency criteria

Article Received: 10 August 2020, Revised: 25 October 2020, Accepted: 18 November 2020

Introduction

Following the emphasis made by Ministry of Education Malaysia (MOE) on cultivating the basic skills of students who attended primary schools, consisting of reading, writing, counting and reasoning, a move to abolish the mid-and final-year examinations for lower primary school pupils has been initiated [1]. This was said to be in accordance with the objectives of School-Based Assessment system (PBS), first implemented back in 2011, which was to assess pupils but not necessarily through the traditional examinations [2]. In place of the traditional examination-based assessment, students will be evaluated through a more objective assessment, known as Classroom-Based Assessment (PBD), which focuses on strengthening the 4M skills of every student that will take place frequently.

This initiative could also be partially attributed to the findings off few research such as [3] which found that students were put an extreme pressure due to the endless waves of examination-based assessment that they were forced to sit for. Besides, outcome from a survey by a national school support service in the United Kingdom (UK), with 1,200 teachers participating, also resulted in rapid growth of anxiety disorders and panic attacks, caused by stresses, which implicated almost 78% primary school pupils in the UK[4]. In striving for the best result while proving their worth to fit in the excellent narrative shown only through examinations, particularly known as SATs, a nationwide standardized tests in this case, it was deduced that students were straining their mental and physical abilities to the point of overexertion which lead to an immense pressure faced by the students.

Therefore, in attaining and sharpening the basic skills of the students while providing a more relaxed environment in learning, the execution of PBD shall take place in a continuous manner throughout the first three years of

learning in primary school. However, despite the distribution of a standard assessment scheme to the teachers in ensuring a better objective evaluation of the pupils as means to measure the success rate of PBD implementation, various factors affecting the objectivity of the evaluation, which could further compromise the desired outcome of the activity, have to be considered. Few known factors influencing the assessment that have been acknowledged in a number of studies are the inability to fully grasp the concept of carrying out formative assessment [5], biased judgement and selective favouritism due to the differing abilities of students in mastering each of the basic skills [6] as well as the lack of interaction between teachers and students which could assist teachers in recognizing the weakness of their students [7].

As attending primary schools was made compulsory for all eligible children, contributing to the huge numbers of primary school pupils, PBD shall also be perceived as a burden in a form of higher workloads to the teachers due to the frequent execution of the assessment. This could also play out in the activity which could result in a less objective and inconsistent assessment by the teachers. As a result, in resolving the inconsistency as well as non-objectivity issues that could play out in PBD, a study is formulated to organize an in-depth investigation on methods of carrying out PBD in a just environment with the aid of information-technology-assisted technique. The scope of study will focus on the reading skills of Year One primary school pupils in terms of literacy fluency while adopting the current speech recognition technique in capturing the rate of fluency. Despite the many techniques developed in recognizing speech, there has yet to be one that could capture while simultaneously assess the fluency of the speech in Malaysia. This is because various studies have been conducted on improving recognition of emotion in speech [8] [9] while

none really looked into evaluating speech fluency based on standard criteria, especially for primary school pupils.

In regards to further elaboration on the intended study, the next section shall discuss on issues concerning the education system as well as reading issues related to children. The follow-up to the section will be the methodology that will be performed throughout the study before the findings from the literature review on speech recognition, of which relevant technique shall be adopted in that of the proposed automated reading fluency evaluation, which will be presented in the section after. The paper will conclude with a conclusion and future recommended study.

Reading Skills and Primary Schoolers

2.1. Issues in the Education System

Education is what people perceived as the one beneficial thing that helps growing a child the most into a better shape but a poorly structured education could also do more harm than good to a child. Despite the essentiality of education, school kills creativity due to the restrictions imposed on the learners. When students are restricted to learning within the boundaries set by their own schools, they are bound to mask their maximum potential, preventing them from unleashing their natural creativity [10].

Furthermore, a poll conducted among 420 members of Association of Teachers and Lecturer (ATL) found that students in their schools do not only have the tendency for self-harm, but there are also a few of them attempting suicide in which 18 of them are primary schoolers[3]. Regardless, such unfortunate event does not only strike the children of Liverpool as it also happens in the other side of the world. The situation within the education system in Singapore was not any different than that of the Liverpool's [11]. Singapore students are equipped with sufficient knowledge that allows them to notch the top spot in Programme for International Assessment (Pisa), a global test which incorporated problem solving, all into science, mathematics and reading evaluation [12].

However, the competitive constraints among the children in Singapore, driven by their parents and schools, have negatively affected their development in terms of behavioural and social skills [11]. In 2015, a staggering number of 27 children and teenagers committed suicide were reported, the highest for more than a decade. In an interview with Howard Tan, he said that these pressures do not only come from the system but most part of them was accounted on the parents who excessively put emphasis on their children's grade achievements [11].

2.2. Reading Problems among Children

Compared to the previous examination-based assessment, PBD works more on evaluating the performance of the pupils in the 4M skills. Even though the pupils may have attended pre-schools where they have already learned the basics of the 4M, but those skills are not something that human, children in particular, could master within a day, especially when it comes to reading. In fact, it requires continuous effort and determination for a 7-year-old to correctly point the spelling for a simple word. However, not

every kid is born with the ability to acquire the same reading skills, regardless of how fast they started learning the alphabets. Furthermore, there are three common difficulties in reading among the pupils which are issues with pronunciations, low comprehension level and reading speed [13].

At times, there are also cases of dyslexic kids, those who suffered from difficulty in translating a written text into a speech. Usually, they find it hard to recognize letters and differentiating the sounds for each letter. As noted by [14], this brain-based condition also affects the reading comprehension of a child due to one's difficulty of getting through a single word. Since they are unable to read the whole sentence, they will barely comprehend what is happening in a story. However, [14] further clarified that having dyslexia does not equate to low intelligence as there are many successful people with dyslexia.

Often being mistaken as dyslexia, Attention-Deficit Hyperactivity Disorder (ADHD) is the most common disorder in childhood where one could hardly pay attention during reading, resulting in their own hyperactivity as one of the symptoms [15]. Although those who suffered from reading issues may be agitated out of frustration or acting up in class just so they could avoid from completing certain tasks, it is actually common for kids to have both ADHD and dyslexia. It was found that a third of kids with attention issues do actually suffer from dyslexia [14].

Reading problems do not only revolve around the previous two brain-based conditions, instead there are more disorders, those of which many are not accustomed to. Auditory processing disorder (APD), a disorder which holds an effect on the child's information processing abilities, often causes one to face difficulty in comprehending what they hear. According to [16], those with APD condition could hardly tell the difference between phonemes, especially letters such as b, d, and p. They also have low literacy skills since reading requires one to be able to connect the pronunciations with the words.

Next is the one disorder where people seem to perceive it the same as dyslexia when it is not, the visual processing disorder [17]. As the name implies, children with this issue will experience hardship in processing every information they come across with. Whenever a child with this issue reads, it may seem hard for him to interpret the words on the book. They also are unable to tell the correct order between letters or even shapes due to their troubled vision and having seen them in double.

2.3. Automatic Speech Recognition

Speech processing could be applied in assisting people throughout their daily routines and few of the most useful applications are speech coding, speech recognition and speaker recognition [18]. Speech recognition, known as Automatic Speech Recognition (ASR), is the process of which a signal that comes from speech is converted to a series of texts by using algorithm-based computer programs. The converted speeches are then distributed into different types of classes which are isolated words, connected words, continuous speech and spontaneous speech [19].

One of the most important steps in recognizing speech is feature extraction. This step is considered the most crucial

part of SRS as its function is to extract critical features from the speech signal that could aid in determining the type of speaker [18][19][20]. For this step, many different techniques could be used in extracting the specific features of a speech depending on the properties of each technique. For example, Mel-Frequency Cestrum Coefficient (MFCCs) method is used due to its robustness in extracting the speech feature[20].

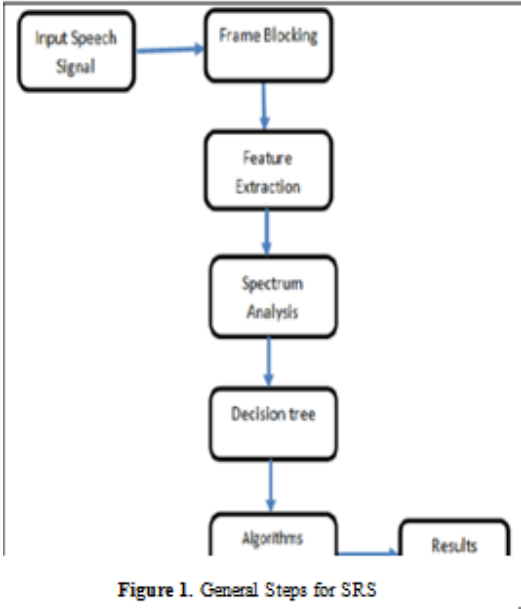


Figure 1. General Steps for SRS

Upon extracting the speech features, spectrum analysis will then be implemented on each block by using either one of the techniques, followed by Fast Fourier Transform (FFT) algorithm and Mel Filter Bank, if one was to use MFCC as the technique [21]. Before applying the right algorithm to obtain the results of the identified speech, decision tree will be performed on each block to single out each phonemes based on their unique attributes [20][21]. The last step which is the implementation of algorithms will be used to refine the implementation of decision tree. To increase the rate of success, different algorithms will be tested until the right one is constructed [20].

The major components of an SRS typically include acoustic front-end, acoustic model, language model, lexicon and search algorithm which acts as the decoder of the inputted speech [22]. The final output will be in the form of a hypothesized word as a result of the transmitted speech signal processing [21]. For any speech signal that is deemed useful in providing information on recognition, acoustic front-end will be the one that acts as a speech signal converter, which converts the speech into their suitable features [18][20][22].

As defined by [18][19][20][21], feature extraction is where the signal of the speech is converted into a suitable form that could later be used at the following steps. It included extracting parameters at times such as the signal magnitude and frequencies power [23]. The most commonly known techniques are Linear Prediction Coding (LPC), Discrete Cosine Transform (DCT), Local Binary Pattern (LBP), Principal Component Analysis (PCA), Linear Predictive Coefficients (LPC), Linear Predictive Cepstral Coefficients (LPCC), Linear Discriminant Analysis (LDA), Weighted Linear Prediction Cepstral Coefficients (WLPCC) [24] and MFCC. However, as a result of a thorough study, it is

evident that MFCC is the one technique that is widely used for speech recognition due to it being a better option even for front-end noise parameterizations [23]. More on feature extraction technique that will be executed in a speech recognition architecture shall be discussed in the later section of this paper.

The outcome of the feature extraction will result in converted feature vectors that will then be fed to the speech decoder where the acoustic models, language models and lexicon are combined altogether. Based on the relevant information from both acoustic and language models, decoder will then proceed to seek for the possible sequence of acoustic class that is compatible with the features sequences [25]. While acoustic modelling is the establishment of statistical representations for the sequence of feature vectors, acoustic models are the file that stores each statistical representations of the varying sounds forming a word which will be labelled as the phoneme.

ResearchMethodology

Aside of determining the methods to be used in collecting data, a research methodology also acts as a systematic plan for doing a research. Hence, in this study, the methodology will be majorly divided into six phases, as shown in Figure 4. The research activities for this study will be divided into six (6) phases with three (3) sub-divisions each. First phase will be the preliminary phase, followed by requirement gathering, conceptual framework development, fuzzy rule-based development, evaluation and lastly, documentation phase which marks the end of the research study.

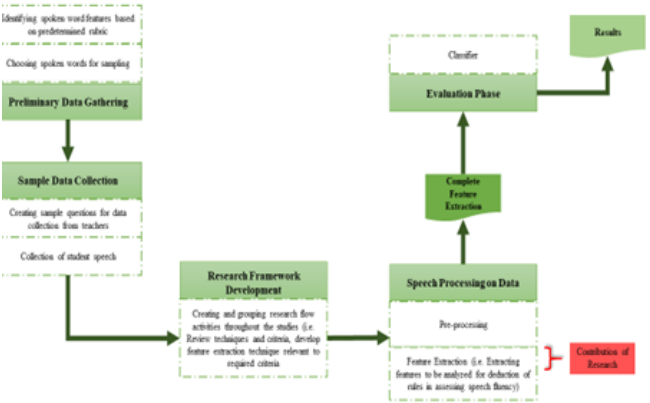


Figure 1. Proposed Research Framework

Since the study is still in the midst of data collection phase, for now, this paper will discuss details pertaining the preliminary data gathering and the components in collecting data for the study.

For this study, the preliminary data gathering phase was made up of two activities which are literature search and actual data collection. Literature search was conducted to identify relevant speech recognition techniques before adapting the necessary components and methods into evaluating reading fluency in this study. By using the likes of IEEE Xplore Digital Library, ACM Digital Library, and ScienceDirect as the online search facilities, more than 89 materials, comprising journals and articles from conference proceeding, have been downloaded using variety of search

terms such as reading fluency, speech recognition and feature extraction in recognizing speech. The articles were then grouped into their respective aspects based on the techniques proposed for the study, case studies that they have referred to as well as the feasibility of the techniques developed. Each group will be analysed for the relevance of incorporating the method and design into that of the proposed speech fluency recognition technique of the study.

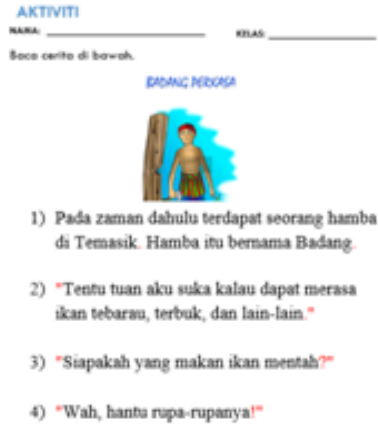


Figure 2. Template of Sample Questions for Students

As for the actual data collection, questions that will be enquired to the teachers involved in teaching and assessing the reading skills of the Year One primary school pupils have been drafted which probes them on the importance of students having acquired the skills of reading at a young age and the significance of technology-assisted technique in assessing students. Figure 3 shows the template in capturing the students’ speech that has been prepared for the purpose of collecting sample data.

Results & Discussions

4.1. Determining the Criteria in Assessing Speech Fluency

As quoted from [26], assessment of student performance is critical for developing effective instructional policy and designing programs responsive to individual students’ needs. Therefore, this is why it is crucial for students to be evaluated in a trustworthy assessment which in turn would be a critical point for each one of them to succeed. Furthermore, [27] also argued on how every teachers or educators should prioritize attaining consistency when it comes to putting each assessed works from the students to a certain standard. For example, teachers who were entrusted to assign large group of students should not be any less subjective in evaluating their students than those with smaller size of students. Given that humans, in their nature itself, have different perceptions as well as differing perceptions, this is where it becomes highly significant for them to unanimously establish a standard set of criteria that will be used as grading guidelines in conducting a thorough and just assessment for the students [28].

Nevertheless, this is no exception, especially for a reading fluency assessment, as different teachers approach the rate of fluency among their students differently. While some may perceive fast reading as fluently excellent read, the other may consider it as mediocre read due to the lack of intonation as well as expression [29][30]. As concluded by [31] upon conducting a consistent global assessment with reasonable outcomes, it came down to the fact that they have already prepared a predetermined set of criteria for all the parties involved in assessing the students. Table 1 indicates the most commonly used criteria in a study that concerns reading fluency or speech recognition.

Table 1. Common Criteria in Assessing Reading Fluency

NO	AUTHOR	CRITERIA							
		PAUSING/SMOOTHNESS	PHRASING	INTONATION/EXPRESSION	STRESS/ACCURACY	INTERGRATION	RAVITY	RATE	GUIDING
1	[32]	/	/	/	/	/	X	/	/
2	[33]	/	/	/	X	X	X	X	X
3	[34]	/	/	/	/	/	X	/	X
4	[35]	/	X	X	/	X	/	X	X
5	[36]	/	X	/	/	X	/	/	X
6	[37]	X	X	/	/	X	/	X	X
7	[38]	/	/	/	X	X	X	X	/
8	[39]	/	/	/	X	X	X	X	/
9	[40]	X	X	X	/	X	X	X	/
10	[29]	/	/	/	X	X	X	X	/
11	[41]	/	/	/	X	X	X	X	/
12	[30]	X	X	/	X	X	X	X	X
13	[42]	/	/	/	X	X	X	X	/
14	[43]	X	/	/	/	X	X	X	/
15	[44]	/	X	/	/	X	X	/	X
16	[45]	X	/	/	/	X	X	X	X
17	[46]	X	/	/	/	X	X	X	/
TOTAL		17	11	15	10	2	3	4	8

It could be assumed that intonation/expression is the one criteria that will be the most looked into when it comes to assessing one’s fluency, as studied by all but [35][40]. Other than that, the fluency of one’s speech will also be determined by their phrasing in reading, considering that 11 out of 15 reviewed case studies opt to use the criteria in their study. Furthermore, it is also essential for one to retain the smoothness [32], [33], [34], [35], [39] as well as the accuracy of their reading to be considered a fluent reader. Moreover, 8 out of 17 studies agreed that the pace of one’s reading should be the indicator of their reading fluency. In conclusion, when evaluating reading fluency of a student, aspects such as intonation, phrasing, smoothness, accuracy and pace of reading should be taken into account for standard reading criteria in an assessment.

4.2. Feature Extraction – The Why’s and How’s

Speech recognition is almost synonymous with every technology-equipped people as it comes in various ways that benefit the mankind. From translating one’s speech to a machine-readable format (e.g. Google Translate) to operating a voice-triggered virtual assistant (e.g. Siri), there have been significant uses of speech recognition. However, in determining the right function of an enhanced speech

recognition, it requires one to select the proper technique, especially feature extraction techniques, for the outcomes they intended for. This is because different techniques hold different abilities, thus the different rate of accuracy in accomplishing certain tasks. Therefore, this section aims to

analyse the relevant techniques to be used, depending on the purposes of the system itself. This will help in determining the most compatible approach that one could have used for in their study, especially that of a reading fluency evaluation.

Table 2.What ASR is for and How Feature Extraction is Performed

Purpose	Author	Technique
To distinguish dysfluent and fluent speech	[47], [48], [49], [50], [51], [52]	Mel-Frequency Cepstral Coefficient (MFCC)
To classify the level of fluency	[53], [54]	Mel-Frequency Cepstral Coefficient (MFCC)
		Convolutional Neural Network (CNN)
To convert text to speech	[55], [56]	Mel-Frequency Cepstral Coefficient (MFCC)
To identify the difference between normal and disordered speech	[57], [58]	Mel-Frequency Cepstral Coefficient (MFCC)
		- Envelope Modulation Spectra (EMS) - Multidirectional Regression (MDR)
To control the movement of mobile robot	[59], [60], [61]	Linear Predictive Analysis (LPC)
		Dynamic Time Wrapping (DTW)
To recognize digit	[59], [62], [63], [64], [65]	Linear Predictive Analysis (LPC)
		Relative Spectral (RASTA) Filtering
		Dynamic Time Wrapping (DTW)
		Mel- Frequency Cepstral Coefficient (MFCC)
To overcome the limitations on software associated with speech recognition	[59]	Linear Predictive Analysis (LPC)
To recognize voice/speech	[66], [67], [59], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83]	Mel- Frequency Cepstral Coefficient (MFCC)
		Zero-Crossing with Peak Amplitude (ZCPA)
		- Mel Frequency Cepstral Coefficients (MFCC) - Principal Component Analysis (PCA)
		- Discrete Cosine Transform (DCT) - Local Binary Pattern (LBP)
		- Mel Frequency Cepstrum Coefficient (MFCC) - Linear discriminant analysis (LDA)

		Dynamic Time Wrapping (DTW)
		First Formant Bandwidth (F1BW)
		• Linear Predictive Coding (LPC) • Linear Predictive Cepstral Coefficients (LPCC) • Weighted Linear Predictive Cepstral Coefficient (WLPCC)
		Relative Spectral (RASTA) Filtering
		Linear Predictive Analysis (LPC)
Throat signal analysis	[84]	Zero-Crossing with Peak Amplitude (ZCPA)
To recognize emotion	[85]	• Linear Predictive Cepstral Coefficients (LPCC) • Mel Frequency Cepstral Coefficients (MFCC)

Studying the comparison between techniques, shown in Table 2, it could be assumed that most studies are interested in exploring the ability of the speech recognition in detecting a voice while only few studied on differentiating fluent speech from the dysfluent ones. Furthermore, all of these studies on fluent and dysfluent speech seem to approach the use of MFCC in their feature extraction techniques. Regardless of the limited number of studies which cover fluency in a broad aspect, it could still be of help to this study in deciding the right techniques in extracting the precise criteria of determining the fluency level of one’s reading.

4.3. Feature Extraction – Extracting Features from Speech Signal as the Criteria of Fluency

There have been various research studies that mainly used MFCC, though not all, as the technique for feature extraction in most study of speech recognition. In detecting speech uttered by users, the information stored in speech signal shall be analysed at the time the words were pronounced, which usually takes place at the classical front end [18][22]. This is called spectral analysis which acts as a parameter in limiting speech signal into that of feature vectors [86]. As MFCC approach uses the base of spectral analysis as the parameters in recognizing speech, it plays a role of a coefficient in representing audio that will be based on what was perceived as the human auditory systems [87]. Furthermore, in obtaining MFCC vectors from each speech frame, FFT of each frame has to be computed before the

magnitude could be collected. This is to enable the execution of a zero-padding to the nearest power of two within speech frame length [86]. Starting with windowing and framing as part of the feature extraction, FFT algorithm will then be applied on the frames of each speech data until the magnitude is found. Upon identifying the magnitude, the signal will be filtered by Mel-Scale Filter Bank, a set of log filter banks that will be arranged according to the Mel Scale [18][22]. Mel Scale is a logarithmic scale that acts as a measurement of the transmitted frequency of a speech tone, which in turn, mimics the magnitude of human’s auditory receptions [22]. As a result of the filtration, the logarithm will be taken from the magnitudes of the log filter banks, enabling the calculation of MFCC that will be done by using discrete cosine transform (DFT) on the coefficients of Mel spectrum.

Table 3. Features Extracted from Different Languages that Match the Criteria of Fluent Reading with the Use of Feature Extraction Techniques

CATEGORY	LANGUAGE	TECHNIQUES	CRITERIA					AUTHOR
			PAUSING / SMOOTHNESS	PHRASING	INTONATION / EXPRESSION	STRESS / ACCURACY	PAGE	
Normal	English	Mel-Frequency Cepstral Coefficient (MFCC)	/	/	/	/	/	[47], [53], [57], [49]
	English & Arabic Numerals (0-9)							[59], [62], [63], [64]
	Marathi							[55]
	Urdu							[67]
	Devanagari							[59]
	Hindi							[48], [56]
	Malay	- Linear Predictive Coding (LPC) - Linear Predictive Cepstral Coefficients (LPCC) - Weighted Linear Predictive Cepstral Coefficient (WLPCC)	X	X	/	/	/	[81]
	English	Convolutional Neural Network (CNN)	X	X	X	X	X	[54]
	Hindi	- Linear Predictive Cepstral Coefficients (LPCC) - Mel Frequency Cepstral Coefficients (MFCC)	/	X	/	/	/	[85]
	Indonesia	- Mel Frequency Cepstral Coefficients (MFCC) - Principal Component Analysis (PCA)	/	/	/	/	/	[80]
	Polish	- Envelope Modulation Spectra (EMS) - Multidirectional Regression (MDR)	/	X	/	X	X	[58]
	English	- Discrete Cosine Transform (DCT) - Local Binary Pattern (LBP)	X	X	/	X	X	[79]
	English	Zero-Crossing with Peak Amplitude (ZCPA)	X	X	/	/	X	[59], [68], [71]
	Korean							[70], [84]
	Chinese							[69]

	Devanagari							[59]
	English	Linear Predictive Analysis (LPC)	X	X	/	/	X	[60]
	Vernacular							[59]
	Devanagari							[66]
	English Numerals							[59]
	English	Dynamic Time Wrapping (DTW)	X	X	/	/	X	[59], [61], [74]
	Arabic							[65], [72]
	English Numeral (1-5)							[73]
	English	Relative Spectral (RASTA) Filtering	X	X	/	X	X	[59], [75], [77], [78]
	Spanish							[76]
	Malay	First Formant Bandwidth (F1BW)	X	X	/	X	/	[83]
	Marathi	- Mel Frequency Cepstrum Coefficient (MFCC) - Linear discriminant analysis (LDA)	/	/	/	/	/	[82]
Disability	English	Mel-Frequency Cepstral Coefficient (MFCC)	/	/	/	/	/	[57], [49], [50], [51]
	Hindi							[48]
	English	- Mel Frequency Cepstrum Coefficient (MFCC) - Linear Predictive Cepstral Coefficients (LPCCs)	/	/	/	/	/	[52]

Based on the review of techniques and features to be extracted as the criteria, it could be assumed that MFCC is the most suitable technique in extracting speech features, be it for normal speech or speech of a person with disabilities such as stutter, dysgraphia, etc. Other than it being frequently used over any other techniques, it also could extract the needed features which match the predetermined criteria in assessing one's fluency [57], [49], [50], [51]. Therefore, as a deduction, by using MFCC for the feature extraction technique, criteria such as pausing, phrasing, intonation, accuracy and pace of reading will be looked into as the features to be extracted in measuring the level of one's fluency.

Conclusion

Concerning the idea of the 4M-based skills assessment that will be implemented for Year One pupils in Malaysia, this research aims to satisfy three research questions at the end

of the study through the research objectives. Firstly, in identifying the criteria needed in evaluating the reading fluency of Year One pupils in Malaysia, an in-depth interview with primary school teachers will be conducted and a reading activity will be organized with Year One pupils. Secondly, in evaluating efficiency of Speech Recognition in capturing speech fluency, existing methods are studied and reviewed into a structured literature review. Thirdly, a proposed technique in extracting speech features to measure the reading fluency of Year One pupils based on the existing feature extraction techniques commonly used in an ASR. While the study may have analysed the techniques and criteria to be complied with during the assessment, data collection phase has yet to take place due to the pending approval. As of now, the study is currently in the phase of collecting sample speech data from Year One pupils in Malaysia.

References

- [1] S. A. Sukaimi, "Move to strengthen lower primary school assessments," 19 December 2018. [Online]. Available: <https://www.nst.com.my/news/nation/2018/12/442191/move-strengthen-lower-primary-school-assessments>. [Accessed 17 January 2020].
- [2] M. Ramasamy and M. A. Patho Rahman, "Maszlee: Abolishing exams for Years One to Three next year aims to restore true spirit of PBS," 31 October 2018. [Online]. Available: <https://www.nst.com.my/news/nation/2018/10/427036/maszlee-abolishing-exams-years-one-three-next-year-aims-restore-true>. [Accessed 17 January 2020].
- [3] J. Burns, "Children as young as six 'stressed' about exams and tests," 4 April 2016. [Online]. Available: <https://www.bbc.com/news/education-35940084>. [Accessed 24 March 2019].
- [4] H. Ward, "Stress of Sats gives children nightmares," 9 July 2018. [Online]. Available: <https://www.tes.com/news/stress-sats-gives-children-nightmares>. [Accessed 7 March 2019].
- [5] A. Z. Khairani and N. T. D. Quyen, "Reviewing the Challenges of Implementing Formative Assessment in Asia: The Need for a Professional Development Program," *Journal of Social Science Studies*, vol. 4, no. 1, pp. 160-177, 7 November 2016.
- [6] P. Afflerbach, *Understanding and Using Reading Assessment*, Newark: International Reading Association, 2017.
- [7] K. Iqbal, Q. Suleman and H. M. Irshadullah, "Understanding the Formative Assessment: A Road Towards Target Achievements," *Journal of Education and Practice*, vol. 8, no. 28, pp. 58-66, 2017.
- [8] M. Neumann and N. T. Vu, "Improving Speech Emotion Recognition with Unsupervised Representation Learning on Unlabeled Speech," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, 2019.
- [9] J. Umamaheswari and A. Akila, "An Enhanced Human Speech Emotion Recognition Using Hybrid of PRNN and KNN," in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Faridabad, India, 2019.
- [10] M. Ross, "The Flaws of Today's Education System," 4 January 2015. [Online]. Available: http://www.teenink.com/opinion/school_college/article/753217/The-Flaws-of-Todays-Education-System/. [Accessed 24 March 2019].
- [11] A. Jelita, "The downsides to Singapore's education system: streaming, stress and suicides," 21 September 2017. [Online]. [Accessed 24 March 2019].
- [12] A. Teng, "Singapore students top OECD global survey in problem solving through teamwork," 21 November 2017. [Online]. Available: <https://www.straitstimes.com/singapore/education/singapore-students-top-oecd-global-survey-in-problem-solving-through-teamwork>. [Accessed 24 March 2019].
- [13] M. Cicerchia, "3 Common Reading Problem for Students," 20 12 2016. [Online]. Available: <https://www.readandspell.com/common-reading-problems>. [Accessed 24 March 2019].
- [14] E. Lapkin, "Understanding Your Child's Trouble with Reading," n.d.. [Online]. Available: <https://www.understood.org/en/learning-attention-issues/child-learning-disabilities/reading-issues/understanding-your-childs-trouble-with-reading>. [Accessed 24 March 2019].
- [15] P. Coppin and S. A. Hockema, "Learning from the information workspace of an information professional with dyslexia and ADHD," in *2009 IEEE Toronto*

- International Conference Science and Technology for Humanity (TIC-STH), Toronto, Canada, 2010.
- [16] M. Y. Boboshko, I. P. Berdnikova, N. V. Maltseva, D. I. Kaplun, V. V. Gulvanskiy and A. S. Voznesenskiy, "Evaluation of human auditory system immunity under central auditory processing disorders," in 2017 XX IEEE International Conference on Soft Computing and Measurements (SCM), St Petersburg, Russia, 2017.
- [17] K. Kelly, "The Difference between Dyslexia and Visual Processing Issues," [Online]. Available: <https://www.understood.org/en/learning-attention-issues/child-learning-disabilities/dyslexia/the-difference-between-dyslexia-and-visual-processing-issues>. [Accessed 24 March 2019].
- [18] S. Karpagavalli and E. Chandra, "A Review on Automatic Speech Recognition Architecture and Approaches," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 9, no. 4, pp. 393-404, 2016.
- [19] F. Mehmood and S. Khan, "An Overview on Speech Recognition System and Comparative Study of Its Approaches," in International Conference on Engineering & Emerging Technologies (ICEET-2014), Lahore, 2014.
- [20] S. S. Naziya and R. R. Deshmukh, "Speech Recognition System - A Review," IOSR Journal of Computer Engineerin (IOSR-JCE), vol. 18, no. 4, pp. 1-9, August 2016.
- [21] P. N. Nasreen, A. C. Kumar and P. A. Nabeel, "Speech Analysis for Automatic Speech Recognition - Survey," in International conference on Computing, Communication and Science, Pune, India, 2016.
- [22] S. Karpagavalli, E. Chandra, R. Deepika, P. Kokila, K. Rani and Usha, "Automatic Speech Recognition: Architecture, Methodologies and Challenges - A Review," International Journal of Advanced Research in Computer Science, vol. 2, no. 6, pp. 326-331, November 2011.
- [23] S. S. Khalid, S. Tanweer, A. Mobin and A. Alam, "A comparative Performance Analysis of LPC and MFCC for Noise Estimation in Speech Recognition Task," International Journal of Electronics Engineering Research, vol. 9, no. 3, pp. 377-390, 2017.
- [24] G. Aggarwal and L. Singh, "Classification of intellectual disability using LPC, LPCC, and WLPCC parameterization techniques," International Journal of Computers and Applications, vol. 41, no. 6, pp. 470-479, 2019.
- [25] R. Kraveva and V. Kravev, "On model architecture for a children's speech recognition dialog system," Blagoevgrad, 2009, pp. 105-110.
- [26] E. J. Kame'enui, L. Fuchs, D. J. Francis, R. Good, R. E. O'Connor, D. C. Simmons, G. Tindal and J. K. Torgesen, "The Adequacy of Tools for Assessing Reading Competence: A Framework and Review," Educational Researcher, vol. 35, no. 4, pp. 3-11, May 2006.
- [27] J. Aper, S. Cuver and D. Hinkle, "Coming to terms with the accountability versus improvement debate in assessment," Higher Education, vol. 20, pp. 471 - 483, 1990.
- [28] J. Balla and P. Boyle, "Assessment of student performance: a framework for improving practice," Assessment and Evaluation in Higher Education, vol. 19, no. 1, pp. 17 - 28, 1994.
- [29] K. Yıldırım, T. Rasinski and D. Kaya, "Fluency and Comprehension of Expository Texts in Turkish Students in Grades Four through Eight," Eğitim ve Bilim , vol. 42, no. 192, pp. 1-12, November 2017.
- [30] P. J. Schwanenflugel and H. Ruston, Becoming a fluent reader: From theory to practice, 2008.
- [31] N. Sakamat and N. E. A. Khalid, "Handwriting Criteria Analysis of Lower Primary School Children," Universiti

- Teknologi MARA Cawangan Terengganu, Malaysia, 2019.
- [32] Lipson and Wixson, "Six Dimensions Fluency Rubric," in *Assessment and Instruction of Reading and Writing Difficulty*, Pearson Education, 2003.
- [33] G. S. Smith and D. D. Paige, "A Study of Interrate Reliability 1 A Study of Reliability Across Multiple Raters When Using the NAEP and MDFS Rubrics to Measure Oral Reading Fluency," *Reading Psychology*, pp. 1-46, January 2019.
- [34] M. Sarris and I. C. Dimakos, "Oral reading fluency and prosody: A preliminary analysis of the Greek language," *Themes in Science & Technology Education*, vol. 8, no. 1, pp. 47-62, October 2015.
- [35] B. L, T. C, V. L, V. E and P. G, "Reading Fluency As a Predictor of School Outcomes across Grades 4 - 9," *Frontiers in Psychology*, vol. 8, no. 200, pp. 1 - 9, February 2017.
- [36] M. Álvarez-Cañizo, P. Suárez-Coalla and F. Cuetos, "The Role of Reading Fluency in Children's Text Comprehension," *Frontiers in Psychology*, vol. 6, no. 1810, pp. 1 - 8, November 2015.
- [37] Z. Elhassan, S. G. Crewther, E. L. Bavin and D. P. Crewther, "Preliminary validation of FastaReada as a measure of reading fluency," *US National Library of Medicine National Institutes of Health*, Bethesda, 2015.
- [38] T. V. Rasinski, "Assessing Reading Fluency," *Pacific Resources for Education and Learning*, Honolulu, Hawaii, 2004.
- [39] C. P. Kor, H. M. Low and L. W. Lee, "Relationship between Oral Reading Fluency and Reading Comprehension among ESL Students," *Gema Online Journal of Language Studies*, vol. 14, no. 3, pp. 19-32, September 2014.
- [40] M. Yildiz, P. K. Uysal, H. Bilge and A. P. Wolters, "Relationships between Turkish Eighth-Grade Students' Oral Reading Efficacy, Reading Comprehension and Achievement Scores on a High-Stakes Achievement Test," *Reading Psychology*, pp. 1-21, July 2019.
- [41] T. Rasinski, "Creating Fluent Readers," *Educational leadership: journal of the Department of Supervision and Curriculum Development*, N.E.A, vol. 61, no. 6, pp. 46-51, March 2004.
- [42] D. Konza, "Fluency," in *Research into practice*, Department of Education and Children's Services, 2011, pp. 1-6.
- [43] B. Education, *Informal Assessments for Fluency Development*, Pelham, New York: Benchmark Education Company, LLC, 2011.
- [44] R. F. Hudson, H. Lane and P. C. Pullen, "Reading fluency assessment and instruction: What, why, and how?," *The Reading Teacher*, vol. 58, no. 8, pp. 702-714, May 2005.
- [45] M. N. Lambert, "A Guide to Reading Fluency and the Assessment of Specific Learning Disabilities in the Individuals with Disabilities Education Improvement Act of 2004," *Wisconsin Department of Public Instruction*, Madison.
- [46] J. Zutell and T. V. Rasinski, "Training teachers to attend to their students' oral reading fluency," *Theory Into Practice*, vol. 30, pp. 211-217, 1991.
- [47] P. Mahesha and D. Vinod, "An Approach for Classification of Dysfluent and Fluent Speech using k-NN and SVM," *International Journal of Computer Science, Engineering and Applications (IJCSEA)*, vol. 2, no. 6, pp. 23-32, December 2012.
- [48] M. Mundada and S. N. Kayte, "Analysis of Speech and its Fluency Disorders," *International Journal of Applied Information Systems (IJAIS)*, vol. 10, no. 4, pp. 10-12, February 2016.
- [49] V. Balaji and G. Sadashivappa, "Waveform Analysis and Feature Extraction from," in *International Conference on Signal Processing and Integrated Networks (SPIN)*, 2019.
- [50] K. N. Vikhyath Narayan and S. P. Meharunnisa, "Detection and Analysis of

- Stuttered Speech," *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, vol. 5, no. 4, pp. 952-955, April 2016.
- [51] A. A. Suryaa and S. M. Vargheseb, "Automatic Speech Recognition System for Stuttering Disabled Persons," *International Journal of Control Theory and Applications*, vol. 10, no. 29, pp. 279-288, 2017.
- [52] S. Oue, R. Marxer and F. Rudzicz, "Automatic dysfluency detection in dysarthric speech using deep belief networks," in *6th Workshop on Speech and Language Processing for Assistive Technologies*, Dresden, Germany, 2015.
- [53] A. Preciado-Grijalva and R. F. Brena, "Speaker Fluency Level Classification using Machine Learning Technique," 2018.
- [54] H. Chung, Y. K. Lee, S. J. Lee and J. G. Park, "Spoken English fluency scoring using convolutional neural networks," in *20th Conference of the Oriental Chapter of the International Coordinating Committee on Speech Databases and Speech I/O Systems and Assessment (O-COCOSDA)*, Seoul, South Korea, 2017.
- [55] S. N. Kayte, M. Mundada and C. Kayte, "Performance Evaluation of Speech Synthesis Techniques for Marathi Language," *International Journal of Computer Applications*, vol. 130, no. 3, pp. 45-50, November 2015.
- [56] S. Kayte, M. Mundada and D. C. Kayte, "Performance Calculation of Speech Synthesis Methods for Hindi language," *IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)*, vol. 5, no. 6, pp. 13-19, 2015.
- [57] M. Mundada, B. Gawali and S. Kayte, "Recognition and Classification Of Speech And Its," *International Journal of Computer Science and Information Technologies (IJCSIT)*, vol. 5, no. 5, pp. 6764-6767, 2014.
- [58] J. Felcyn and A. B. Cavalcante, "Automatic differentiation between normal and disordered speech," in *Language & Technology Conference 2015*, Poznań, 2015.
- [59] P. K. Kurzekar, R. R. Deshmukh, V. B. W. and P. P. Shrishrimal, "A Comparative Study of Feature Extraction Techniques for Speech Recognition System," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 12, pp. 18006-18016, December 2014.
- [60] R. Singh and P. Rao, "Spectral Subtraction Speech Enhancement with RASTA Filtering," in *Proceeding of National Conference on Communications (NCC)*, Kanpur, India, 2007.
- [61] S. Verma, T. Gulati and R. Lamba, "Recognizing Voice For Numerics Using MFCC and DTW," *International Journal of Application or Innovation in Engineering & Management*, vol. 2, no. 5, pp. 127-130, 2013.
- [62] F. Itakura, "Minimum Prediction Residual Principle Applied to Speech Recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 23, no. 1, pp. 67-72, February 1975.
- [63] G. Vyas and B. Kumari, "Speaker Recognition System Based On MFCC and DCT," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 2, no. 5, pp. 145-148, June 2013.
- [64] S. Iqbal, T. Mahboob and M. S. H. Khiyal, "Voice Recognition using HMM with MFCC for Secure ATM," *International Journal of Computer Science Issues (IJCSI)*, vol. 8, no. 6, pp. 297-303, November 2011.
- [65] N. Desai, K. Dhameliya and V. Desai, "Recognizing voice commands for robot using MFCC and DTW," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, no. 5, pp. 6456-6459, May 2014.
- [66] Thieng and S. Wijoyo, "Speech Recognition Using Linear Predictive

- Coding and Artificial Neural Network for Controlling Movement of Mobile Robot," in International Conference on Information and Electronics Engineering IPCSIT, Singapore, 2011.
- [67] A. S. Nigade and J. S. Chitode, "Throat Microphone Signals for Isolated Word Recognition Using LPC," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 8, pp. 906-913, August 2012.
- [68] S. Gaikwad, B. Gawali and P. Yannawar, "Performance Analysis of MFCC & DTW for Isolated Arabic Digit," International Journal of Advanced Research in Computer Science, vol. 2, no. 1, pp. 513-518, 2011.
- [69] J. Manikandan and B. Venkataramani, "Design of a real time automatic speech recognition system using Modified One Against All SVM classifier," Microprocessors and Microsystems, vol. 35, no. 6, pp. 568-578, August 2011.
- [70] X. Zhang and W. Liang, "A Robust Speech Recognition Based on the Feature of Weighting Combination ZCPA," in First International Conference on Innovative Computing, Information and Control - Volume I (ICICIC'06), Beijing, China, 2006.
- [71] Doh-Suk-Kim, J.-H. Jeong, S.-Y. Lee and R. M. Kil, "Comparative Evaluations of Several Front Ends for Robust Speech," Advance Institute of Science and Technology Korea, 1997.
- [72] S. Haque, R. Togneri and A. Zaknich, "Zero Crossings with Peak Amplitudes and Perceptual Features for Robust Speech," School of Electrical, Electronic and Computer Engineering University of Western Australia, 2012.
- [73] K. A. Darabkh, A. F. Khalifeh, B. A. Bathech and S. W. Sabah, "Efficient DTW-Based Speech Recognition System for Isolated Words of Arabic Language," International Scholarly and Scientific Research & Innovation, vol. 7, pp. 106-113, May 2013.
- [74] Z. Hachkar, A. Farchi, B. Mounir and J. El Abbadi, "A Comparison of DHMM and DTW for Isolated Digits Recognition System of Arabic Language," International Journal on Computer Science and Engineering (IJCE), vol. 3, no. 3, pp. 1002-1008, 2011.
- [75] M. D. Wachter, M. Matton, K. Demuynck, P. Wambacq, R. Cools and D. V. Compernelle, "Template Based Continuous Speech Recognition," IEEE Transactions On Audio, Speech & Language Processing, pp. 898-901, 2007.
- [76] H. Rahali, Z. Hajaiej and N. Ellouze, "Robust Features for Speech Recognition using Temporal Filtering Technique in the presence of Impulsive Noise," I.J. Image, Graphics and Signal Processing, vol. 11, pp. 17-24, 2014.
- [77] P. P. Marsal, S. P. Font, A. Hagen, H. Bourlard and C. Nadeu, "Comparison And Combination Of Rasta-Plp And Ff Features In A Hybrid Hmm/Mlp Speech Recognition System," IEEE Transactions on Speech and Audio Processing, vol. 13, no. 1, pp. 14-22, 2004.
- [78] M. Holmberg, D. Gelbart and W. Hemmert, "Automatic Speech Recognition with an Adaptation Model Motivated by Auditory Processing," IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 1, pp. 43-49, 19 December 2005.
- [79] J. Abhilash and G. N. Rathna, "Visual Speech Recognition for Isolated Digits using Discrete Cosine Transform (DCT) and Local Binary Pattern Features," in Global Conference on Signal and Information Processing (GlobalSIP), 2017.
- [80] A. Winursito, R. Hidayat and A. Bejo, "Improvement of MFCC Feature Extraction Accuracy using PCA in Indonesian Speech Recognition," in International Conference on Information and Communications Technology (ICOIACT), 2018.

- [81] Y. F. Chong, H. M, S. Yaacob and A. Ah, "Malay speech recognition in normal and noise condition," in International Colloquium on Signal Processing and Its Applications (CSPA), 2012.
- [82] S. Gaikwad, B. Gawali, P. Yannawar and S. Mehrotra, "Feature Extraction using Fusion MFCC for Continuous Marathi Speech Recognition," in 2011 Annual IEEE India Conference, Hyderabad, India, 2011.
- [83] S. A. M. Yusof, F. Siraj, S. Yaacob and M. P. Paulraj, "Improved Malay Vowel Feature Extraction Method Based on First and Second Formants," in Second International Conference on Computational Intelligence, Modelling and Simulation (CIMSIM), 2010.
- [84] S. Haque, R. Togneri and A. Zaknich, "Zero-Crossings with Adaptation for Automatic Speech Recognition," in Proceeding of the 11th Australian International Conference on Speech Science & Technology, Auckland, New Zealand, 2006.
- [85] B. Sweetta and D. Amita, "Emotional Hindi Speech: Feature Extraction and Classification," in 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2015.
- [86] A. Nabeel, N. Nasreen and C. Kumar, "Speech Analysis for Automatic Speech Recognition - Survey," in International conference on Computing, Communication and Science (ICRTCCS), Pune, India, 2016.
- [87] P. L. Borde, A. Varpe, R. Manza and P. Yannawar, "Recognition of Isolated Words using Zernike and MFCC features for Audio Visual Speech Recognition," International Journal of Speech Technology, vol. 17, no. 2, pp. 1 - 11, July 2014.