BENTHAM OPEN

CrossMark

The Open Biomedical Engineering Journal

Content list available at: www.benthamopen.com/TOBEJ/

RESEARCH ARTICLE

A Novel Application System of Assessing the Pronunciation Differences Between Chinese Children and Adults

Xiaoyang Zhang[1], Lei Xue[1,*], Zhi Zhang[1] and Yiwen Zhang[2]

[1]*School of Communication and Information Engineering, Shanghai University, Shanghai, 200000, P.R. China*
[2]*Shanghai Children's Medical Center, Shanghai, 200000, P.R. China*

Abstract:

Background:

Health problems about children have been attracting much attention of parents and even the whole society all the time, among which, child-language development is a hot research topic. The experts and scholars have studied and found that the guardians taking appropriate intervention in children at the early stage can promote children's language and cognitive ability development effectively, and carry out analysis of quantity. The intervention of Artificial Intelligence Technology has effect on the autistic spectrum disorders of children obviously.

Objective and Methods:

This paper presents a speech signal analysis system for children, with preprocessing of the speaker speech signal, subsequent calculation of the number in the speech of guardians and children, and some other characteristic parameters or indicators (*e.g* cognizable syllable number, the continuity of the language).

Results:

With these quantitative analysis tool and parameters, we can evaluate and analyze the quality of children's language and cognitive ability objectively and quantitatively to provide the basis for decision-making criteria for parents. Thereby, they can adopt appropriate measures for children to promote the development of children's language and cognitive status.

Conclusion:

In this paper, according to the existing study of children's language development, we put forward several indicators in the process of automatic measurement for language development which influence the formation of children's language. From the experimental results we can see that after the pretreatment (including signal enhancement, speech activity detection), both divergence algorithm calculation results and the later words count are quite satisfactory compared with the actual situation.

Keywords: Assessment, Cognitive ability, Decision-making, Language words, Speech of children, Speech signal processing.

INTRODUCTION

In the recent years, with the understanding of children's language development law, workers who pay attention to children's language development are increasingly urgent to need artificial intelligence technology to automatically monitor and identify children's language development situation.

More and more families are paying close attention to the healthy growth of children. Studies of children's language

* Address correspondence to this author at the School of Communication and Information Engineering, Shanghai University, 200000, P.R. China; Tel: 86-136-3661-7211; Email: 16301098@163.com.

and cognitive development have turned up much. T. R. Risley *et al.* [1] found in the language development of children: (i) the variation in children's IQ and speech language abilities is relative to the number of words which their parents speak to children, (ii) children's academic successes at age of nine to ten are attributable to the number of conversation which they hear from birth to three. And (iii) the parents of advanced children talk significantly more to their children than the parents of children who are not as stand out. Also related studies show that a positive response behavior to children throughout the early child development, which have a positive effect on children's language, cognition, social and emotional behavior [2, 3]. In [4], research suggests that the ability of children to use language and understand the meaning of spoken and written words is related to achievement in reading, writing and spelling later.

Studies also show that parents from low-income families have lack of communication with their children, so the children are less frequently than peers in both speaking and reading, which prevents their ability to develop literacy and language skills. This results in widening achievement gap or "language gap".

In addition, Autism Spectrum Disorder (ASD) has gained more and more attentions in the recent years, and quantitative measurement of language ability plays a very important role in early diagnosis and intervention [5, 6].

At present, there are lots of researches on speech signal analyzing, processing and other aspects. Related technologies have been much more mature. For quantitatively evaluation of the difference between children and adult pronunciation, some researchers have put forward the evaluation system of children's language development which has been applied in foreign countries. However, it is designed for children whose native language is English [7].

There are many differences in environment, language and so on between English and Chinese. For example, English word is composed of vowel and consonant, while Chinese character is composed of initials and finals. Therefore, we cannot directly use analytical methods and parameters when the intelligent segmentation of Chinese words. And in the system, there are few parameters which can evaluate the difference of children and adults pronunciation to reflect the difference between children and adult in pronunciation quality well. Thus, it cannot be applied on Chinese children as their mother tongue is Chinese [8].

There have been some research results about the amount of interaction between adults and children, but for the research and evaluation of children's pronunciation quality were still insufficient. Meanwhile, there is no research or application evaluation of pronunciation differences between children and adults in our country, for now we use a way of scale which is subjective, qualitative to judge the level of children's pronunciation, it can't analysis the status of pronunciation level objectively and quantitatively.

In this paper, based on the early research in characteristics and process of language interaction between parents and children, and combined with the speech signal processing and analysis, we have designed and completed the system for the evaluation of pronunciation parameters and differences between children and adults whose mother tongue is Chinese.

SYSTEM OVERVIEW

Overall Design of System

The overall design idea of the speech signal processing system block diagram proposed in this paper is shown in Fig. (**1**).

First we acquire the speech signal of the target child (study object) in the actual environment (mainly refers to the family environment). The speech signal is collected by special digital voice recording system. Then we extracted the feature vectors which are easy to distinguish between different speakers or the characteristics of the signal source, to facilitate speech signal classification later. In [9], we have achieved separating and clustering the speech signal for the part of children and adults, and the accuracy rate reached more than 80%. But the subsequent analyses of the children's pronunciation quality is rare, and we have not evaluated the quality of pronunciation.

Therefore, this paper studies and presents several methods of speech signal processing and analyzing for the children. The follow-up section of this paper will introduce each module in detail.
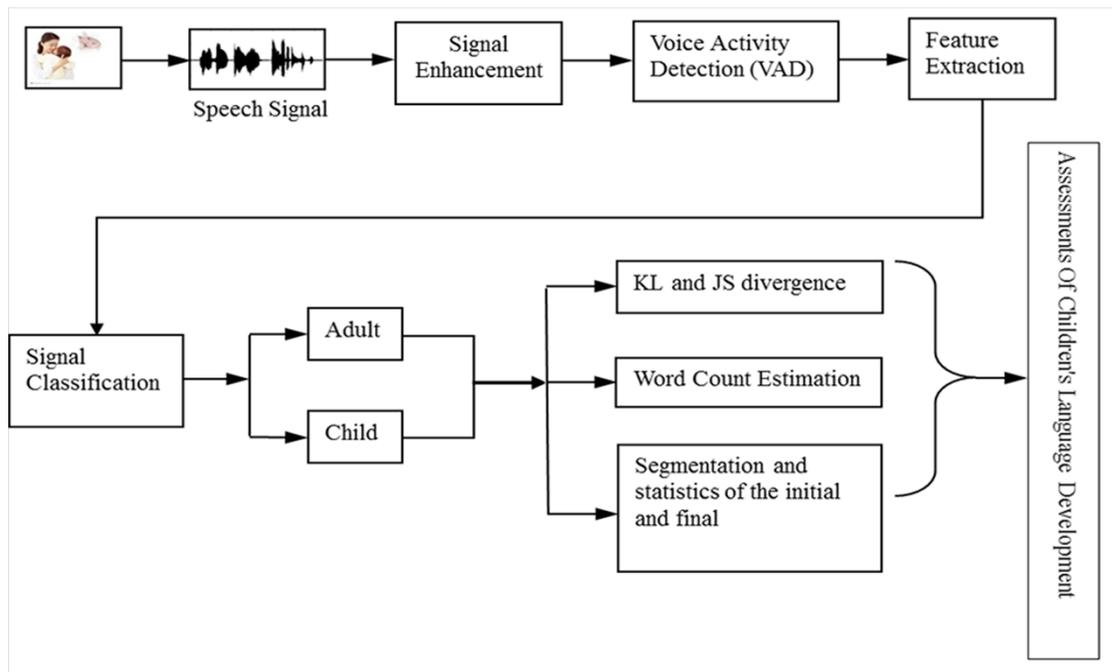
Fig. (1). The system function block diagram of the overall design. After signal enhancement, voice activity detection and feature extraction, we classify the signals, get adult speech and child speech respectively, thus we can calculate the speech parameters and assess the child's speech development.

Automatic Child Vocalization Assessment

　　Automatic Child Vocalization Assessment (ACVA) is the most important feature in this processing system. In order to verify the feasibility and effectiveness of the system, we choose the speech data source which is applied by the language department of Shanghai Children's Medical Center. We record the related family daily conversation between children and adults with a small light-weight digital recorder. The data source of speech are children's all day activity time including they have conversation with their parents and other relatives, bedtime stories, and other language exchange time, as long as 16 hours approximately. The data sampling frequency is 16 kHz, the precision is 16bit. We choose the data source is general, namely family selection is random, and the voice is also under the natural scenes of life. And it has been recognized by the ethics committee.

KL AND JS DIVERGENCE OF THE SPEECH SIGNAL

KL and JS Divergence

　　In probability theory and information theory, Kullback-Leibler divergence, which is also called relative entropy, is a method to describe the difference between P and Q (which are two different probability distributions), and used to measure the distance between two probability distributions. Jensen-Shannon divergence is also a very popular method which measures the similarity between two probability distributions. It is also known as information radius (IRad) and total average divergences. It is based on the Kullback-Leibler divergence, and some other well-known and divergences. At the same time, it is symmetrical and limited [10]. Their formulas are as follows:

$$\text{KL divergence: } D_{KL}(Q \parallel P) = \sum Q(x) \log \frac{Q(x)}{P(x)} \quad x \in X \tag{1}$$

$$\text{JS divergence: } D_{JS}(P \parallel Q) = \frac{1}{2}[D_{KL}(P \parallel M) + D_{KL}(Q \parallel M)] \quad M = \frac{1}{2}(P + Q) \tag{2}$$

　　We used the feature of the KL and JS divergence to judge the similarity between two probability distributions, and assessed the distance between two different speech signals after appropriate mathematical transformation. Make it applied to judge the language pronunciation differences between children and adults, then we can study the situation of children's language development.

According to the characteristics of the speech signal, we used two angles of signal characteristic: ① Time-domain waveform, ② Energy. The processing algorithm of experimental data is as follows:

1. Select the processing section of the experimental data;
2. ABS;
   The data of the speech file exists above and below the x-axis after reading by Matlab, and it means that the data has both positive and negative values. According to the formula of the KL divergence, the effective values range from 0 to infinity, so we take the ABS value.
3. Normalization.
   Two sets of data which are processing must satisfy a certain probability distribution, so the probability distribution must meet the condition of sum is 1, thus it needs to be normalized [11].

Speech Signal Processing

We picked up two speech signals which have same length from the data source, they belong to children's (the following expressed in P) and his father's (the following expressed in Q) respectively, after a series of pretreatment, like framing the signal and calculating the short-time energy, we get real research data we want, and Fig. (**2**) shows the result data.
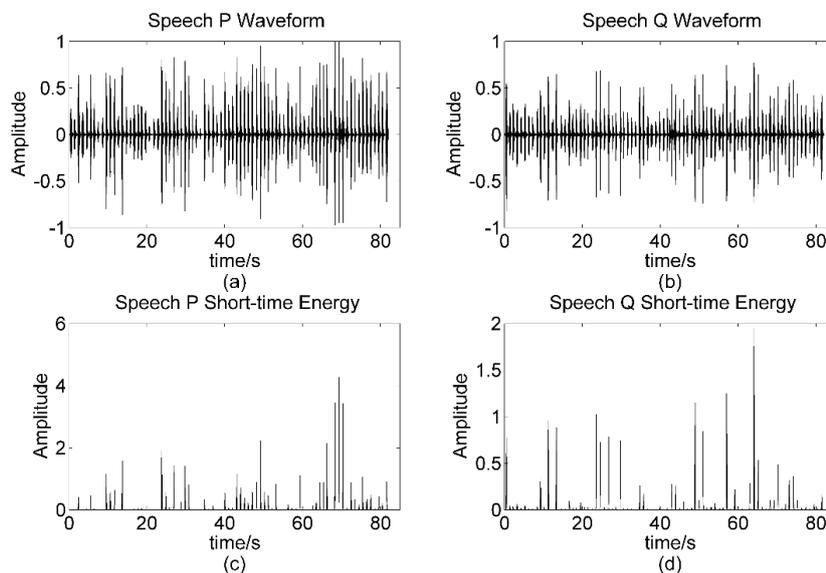

Fig. (2). Time-domain waveform and short-time energy for children and adults. Two parameters of children and adults speech.

Table **1** shows the divergences between two selected recording files, we calculate the KL and JS divergence at the angle of the time domain and short-term energy concurrently when we finished separating them into adult and child.

Table 1. The KL and JS divergence of speech signal between the child and adult.

File Name	Time Domain				Short-Time Energy			
	P→Q		Q→P		P→Q		Q→P	
	KL	JS	KL	JS	KL	JS	KL	JS
20141128.wav	3.9625	0.6606	4.4470	0.6606	13.9625	1.1901	16.4470	1.1901
20150128.wav	3.0260	0.5722	3.9296	0.5722	13.2546	0.9431	15.6041	0.9431

From the results shown in Table **1**, we found that children's language level can be assessed effectively and quantitatively with KL and JS divergence whether from the point of view time domain or short-time energy, and the subsequent series of experiments and test. The smaller the distance, the smaller the speech level gap. It proved that the parameters can be applied to automatically assess the speech signal of children's.

WORD COUNT

As indicated by Hart and Risley [1], adult word count (AWC) is one of the most important factors contributing to child language environment. As one tries to speak more to child, the topic and content tend to be diversified, and so is the vocabulary. We can observe the language development of children from angle of statistics through counting the number of words in dialogues [12].

Short-time Energy Threshold Method

The word count algorithm that this paper is using is based on speech signal short-time energy. We labeled and divided the speech signal, and count them at the same time through setting appropriate threshold values. The size of the threshold is decided by the result after comparison between the real word count among dialogues and the number from the algorithm. We continue to adjust its size to achieve the optimal threshold, and make the algorithm error rate drops [13].

Specifically, the formula for short-time energy of the speech signal is determined as:

$$E_n = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]^2 \qquad (3)$$

Where x (m) is a speech sequential, w (n) is a window function for the corresponding frame. The short-time energy reflects the rule of the amplitude or energy of speech changes over time. In general, the amplitude of the voiced sounds is larger than that under unvoiced sounds, and energy of the voiced sounds is much greater than that under unvoiced sounds. Meanwhile, in high SNR environments, we can use the short-time energy to determine audio sound and silent part, initial and final, even the word boundary, *etc* [14].

The procedure of short-time energy threshold algorithm for speech word count is described below:

1. Pre-processing the speech signal, filter to remove 50Hz frequency interference and background noise;
2. Framing the speech signal, and calculate the short-time energy of speech signal for each frame;
3. Set the value of short-time energy threshold: $E_t = p * max(E_n)$, P-value is an empirical coefficient values, generally the value range is [0.2, 0.4];
4. Record the first frame subscript e_1 and the last frame subscript e_2 which meets the condition $E_n > E_t$
5. Turn the frame subscript into point in time, and according to the value of $e_2 - e_1$ to judge whether the speech segment is true voice or noise, if speech length is too short, we think that is noise, we normally think that 10ms is the shortest length of voice section.

Double Threshold Method Based on Short-time Energy and the Zero-crossing Rate

We also choose another method to carry out the contrast experiment: Double threshold method which is based on short-time energy, and the zero-crossing rate. It is a typical method of endpoint detection. It is combined with two characteristic parameters which are the short-time energy and the zero-crossing rate to detect endpoint. In the practical application, we normally use energy to detect voiced sounds, and use zero-crossing rate to detect unvoiced sounds. We can achieve reliable voice detection in cooperation with using two parameters simultaneously.

The signal will pass through zero when its amplitude value from positive to negative or from negative to positive, it is known as zero-crossing, and the zero-crossing rate of the signal refers to the number of time that the signal cross zero in unit time. The short-time zero crossing rate of speech signals are defined as follows:

$$Z_n = \sum_{m=-\infty}^{\infty} |sgn[x(n)] - sgn[x(n-1)]| \, w(n-m) \qquad (4)$$

Where x (m) is a speech sequential, and *sgn*[] is the symbolic function:

$$sgn(n) = \begin{cases} 1, & x(n) \geq 0 \\ 0, & x(n) < 0 \end{cases} \qquad (5)$$

$$w(n) = \begin{cases} \dfrac{1}{2N}, & 0 \leq n \leq N - 1 \\ 0, & else \end{cases} \tag{6}$$

Zero-crossing rate is time domain characteristics of acoustic parameters, and reflects the spectrum characteristics of speech indirectly, and it can be used to determine the speech whether is unvoiced or voiced. Short-time zero-crossing rate represents the number of speech signal waveform in a frame speech passes through the horizontal axis (zero level). Zero-crossing analysis is one of the most simple time-domain analyses in speech signal analysis. For continuous speech signal, crossing zero line means that the time-domain waveform across the timeline. But for the discrete signal, if adjacent sample values change sign, we call it zero-crossing. In one word, the zero rate is the number of sample changes sign.

The process of speech signal word count algorithm which is based on double threshold method is described as follows:

1. Read in the speech files, and pre-processing the speech signal, filter to remove 50Hz frequency interference and background noise;
2. Framing the speech signal;
3. Calculate the short-time energy and short-time zero-crossing rate for each frame speech;
4. Finding the rising edge of short-time energy and the falling edge of short-time zero-crossing;
5. Picking up the frames which both have the rising edge of short-time energy and the falling edge of short-time zero-crossing and recorded as syllable characteristics, then calculate the number of the speaking words in the speech.

Word Count Estimation

We used two methods which are Short-time energy threshold method and Double threshold method based on short-time energy and the zero-crossing rate to carry out contrast experiment in counting words of the speech. We used four group thresholds comparative tests in the short-term energy threshold method.

Table 2. Experimental results for word count.

File Name	Real Word Count	Short-time Energy Threshold Method			Double Threshold Method Based on Short-time Energy and the Zero-crossing Rate	
		Threshold	Calculate Word Count	Error	Calculate Word Count	Error
1	496	0.20	455	8.26%	469	5.44%
		0.25	479	3.42%		
		0.30	476	4.03%		
		0.35	460	7.26%		
2	557	0.20	581	4.31%	534	4.13%
		0.25	582	4.48%		
		0.30	547	1.80%		
		0.35	500	10.23%		
3	94	0.20	105	11.70%	89	5.32%
		0.25	102	8.51%		
		0.30	100	6.38%		
		0.35	72	23.40%		

The experimental results are shown in Table **2**, the error results of three groups of experimental data are lower than other threshold, and the effect is better in the threshold of 0.3 when we are choosing the short-time energy threshold method. When compared with another method, in comparison to the experiment of first two data sets, the error in short-time energy threshold method is lower. But in the third set of data, the result of double threshold detection method is low, but the difference is not great. Due to the influence of different data length, the data length is bigger which is used in the first two experiments, and finally have an effect on the results of the experiment.

SEGMENTATION AND STATISTICS OF THE INITIAL AND FINAL

The number of species of the initial and final in the mandarin can reflect the level of pronunciation, the larger

species number, the higher lever of speech. On the other hand, the smaller the number of species, the lower the level of speech. Therefore, we can assess children's language development from the angle of the size of species of the initial and final.

The Phenomenon of Coarticulation

The cue of the initial and final is not always within their respective segment in mandarin, and it may cross the border between the two. It means that initial may carry final's information inside. Final also may carry initial's information inside, such as consonants juncture. If the final is count from vowel starts, then juncture is the consonant cue in finals. For some consonants, such as unaspirated plosive: [b], [d], vowel juncture is right the main cue for distinguish them.

But there is a phenomenon that certain cue has a coarticulation, the most typical is the way to achieve sound medium, such as the starting phoneme in compound finals: [i], [u], [y], medial is instable, The implementation of medial is very different under the influence of different initials. We can get the following arrangement according to how many cue is under initials: (many) aspirated affricate-aspirated plosives-fricative-lateral-unaspirated affricate-unaspirated plosive-nasal-zero initial (few).

The bands which signal energy of initials and finals concentrated can be distinguished. Initials signal is shorter than the finals, and the spectrum of finals is more stable than that of the initials'.

Statistical Theory of the Initials and Finals in Speech Signal

It is the "short-time analysis" that throughout the whole process of the speech signal analysis. The characteristics and parameters which represent its essential characters are all changed over time in the speech signal, so it is a non-stationary process, and it cannot use digital signal processing technique which is processing steady signal to analyze and process. However, different speech is produced by the response of sound track which is shaped by human mouth muscles movement, and the frequency of this kind of mouth muscle movement is very low relate to the speech's. Therefore, although the speech signal has time-varying characteristics, the characteristics remain unchanged basically within a short time (It is generally believed in 10 to 30ms). Thus, it can be regarded as a quasi-steady state process, and it means that the speech signal has a short-term stability [15]. As a result, any analysis and processing for speech signal must be established on the basis of the "short-time", namely short-time analysis. The speech signal is divided into paragraphs to analyze its characteristic parameters, and each paragraph is called a frame which generally long for 10 to 30ms. Therefore, in terms of the whole speech signal, the analysis result is the characteristic parameters of time series which are composed of each frame's feature parameters. The time domain analysis of speech signal aims to analyze and extract the time domain parameters. While carrying out speech analysis, the first contact and the most intuitive is its time-domain waveform. The speech signal is a time domain signal. The time-domain analysis is the first to use, and it is the most widely used method of analysis [16].
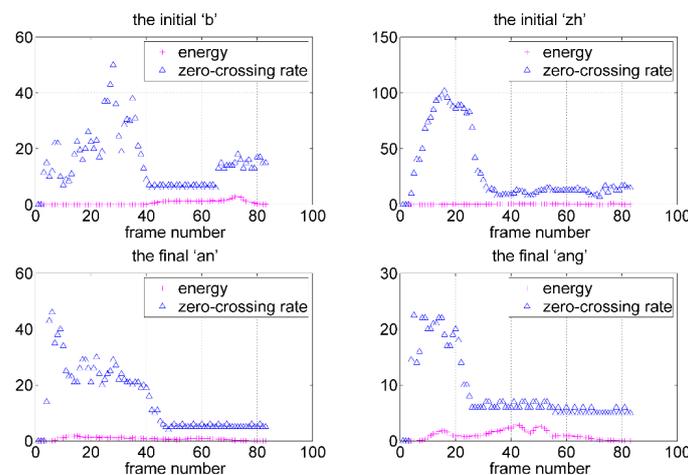

Fig. (3). Comparision between short-time energy and short-time zero-crossing rate of several initials and finals. In Chinese, 'b' and 'zh' are initials, 'an' and 'ang' are finals. It infers that it has a lower zero-crossing rate for finals, a higher rate for initials.

For speech signals, high frequency has high average zero-crossing rate, low frequency has low average zero-crossing rate, and we can refer that it has a lower zero-crossing rate for finals, a higher rate for initials. That, certainly, is just relatively in high or low, and have no precise numerical relationship. Fig. (**3**) shows several different vowel sound, short-time energy and short-time zero-crossing rate. From above observation and analysis, we can refer that: for short-time energy, lower in initial and higher in final; for short-time zero-crossing rate: higher in initial and lower in final. It can be regarded as a quasi-steady process when we are handling characteristics of the initials and finals in speech signal, and it means that the speech signal has short-time stability. In mandarin, the cue of the initial and final is not always in their respective segment, and it may cross the border between the two. It means that initial may carry final's information [17]. The bands of initials and finals of signal energy concentrated can be distinguished, the length of the initial signal is shorter than final, and the frequency spectrum of final is more stable than initial. We can easily distinguish initial and final by analyzing their feature attributes.

Segmentation and Statistics

When it comes to the part of the experimental simulation which is the segmentation and statistics for the initial and final in children's speech, we selected a child's voice file for the separation, and the result is shown in Fig. (**4**). The figure above is the original speech time-domain signal, and between the green and red two vertical lines in figure below is a representative of the word, the black line represents the dividing line of the initial and final, the light blue represents the signal energy after standardized. Followed along the timeline are: 'L', 'an', 'T', 'ian', 'B', 'ai', 'Y', 'un'.
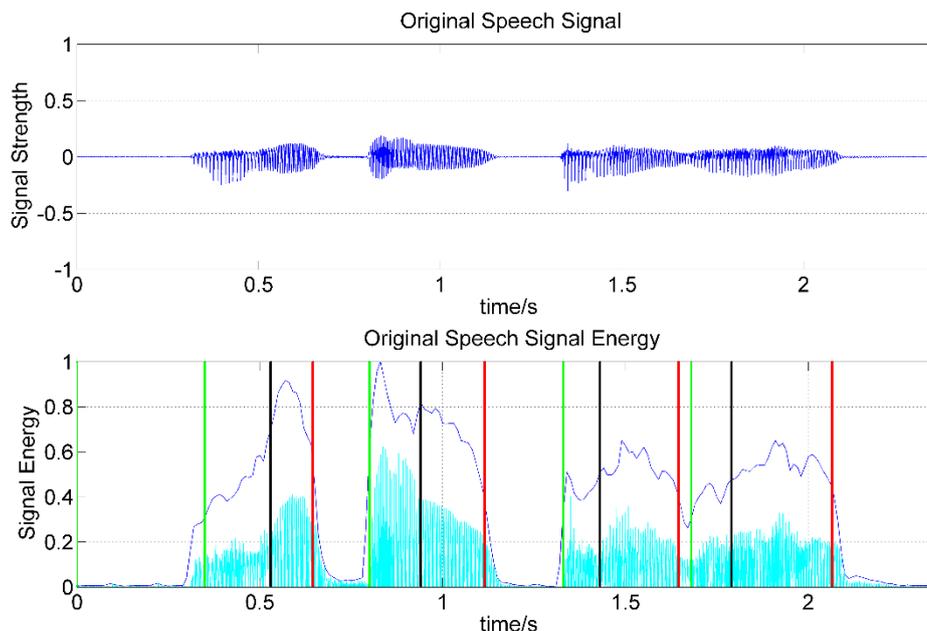

Fig. (4). Separation of the initial and final for child's speech signal. The black lines represent the divider between the final and initial, the part between the green line and the next red line represents a word.

The results in Fig. (**4**) show that the division of children's speech is accurate, it is based on the word segmentation, and then we carry out left segmentation. Meanwhile, it provides the prerequisite for the subsequent classification [18].

CONCLUSION

In this paper, according to the existing study of children's language development, we put forward several indicators in the process of automatic measurement for language development which influence the formation of children's language. The parameters and figures can show the quality of children's language development, and it changes the form of the traditional manual scale which is for detecting children's language development. We provide a way to study and evaluate the condition of children's language development quantitatively and objectively, we can also think it is a potential research direction for us in the future. And it can detect such flaws in the aspect of language development as soon as possible for early treatment, to improve children's language development and language cognitive ability.

From the experimental results we can infer that, after the pretreatment (including signal enhancement, speech activity detection), whether divergence algorithm calculation results or the later words count, and the results are quite satisfactory as compared with the actual situation.

However, if we wish to make a big breakthrough, we have a lot of related work to do, such as improvement of the accuracy of signal separation, increasing the usage of data source, and setting up of experimental control group.

In the future, we will plan to do much more evaluation works in our system to realize truly automation of children's speech signal processing. The system will save energy and time of guardians and health care workers, and help children to have a healthy development of speech and cognitive ability.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   T.R. Risley, B. Hart, and L. Bloo, "Meaningful differences in the everyday experience of young american children", *Child Lang.,* vol. 165, no. 4, p. 282, 1995.

[2]   G. Dawson, and J. Osterling, *Early intervention in autism in the effectiveness of early intervention.,* Brookes: Baltimore, 1997.

[3]   M. Siller, and M. Sigman, "Modeling longitudinal change in the language abilities of children with autism: parent behaviors and child characteristics as predictors of change", *Dev. Psychol.,* vol. 44, no. 6, pp. 1691-1704, 2008.
[http://dx.doi.org/10.1037/a0013771] [PMID: 18999331]

[4]   S.F. Warren, and N.C. Brady, "The role of maternal responsivity in the development of children with intellectual disabilities", *Ment. Retard. Dev. Disabil. Res. Rev.,* vol. 13, no. 4, pp. 330-338, 2007.
[http://dx.doi.org/10.1002/mrdd.20177] [PMID: 17979201]

[5]   National Institute for Literacy, "Learning to Talk and Listen: An Oral Language Resource for Early Childhood Caregivers", 2009. Available at: https://lincs.ed.gov/publications/pdf/LearningtoTalkandListen.pdf

[6]   S.E. Bryson, S.J. Rogers, and E. Fombonne, "Autism spectrum disorders: early detection, intervention, education, and psychopharmacological management", *Can. J. Psychiatry,* vol. 48, no. 8, pp. 506-516, 2003.
[PMID: 14574826]

[7]   Lena Research Foundation. Available at: http://www.lenafoundation.org.

[8]   D. Xu, U. Yapanel, S. Gray, J. Gilkerson, J. Richards, and J. Hansen, "Signal processing for young child speech language development", In: *1ˢᵗ Workshop on Child, Computer and Interaction,* Chania, Crete, Greece, 2008.

[9]   L. Xue, Z. Zhang, X. Zhang, and Y. Zhang, "Research and implementation of children's speech signal processing system", *Open Biomed. Eng. J.,* vol. 9, 2014.

[10]   D.R. Anderson, and G.C. White, "Kullback-Leibler information as a basis for strong inference in ecological studies", *Wildl. Res.,* vol. 28, no. 2, pp. 111-119, 2001.
[http://dx.doi.org/10.1071/WR99107]

[11]   D.R. Anderson, and G.C. White, "Kullback-Leibler information in resolving natural resource conflicts when definitive data exist", *Wildl. Soc. Bull.,* vol. 29, no. 4, pp. 1260-1270, 2001.

[12]   S.F. Warren, J. Gilkerson, J.A. Richards, D.K. Oller, D. Xu, U. Yapanel, and S. Gray, "What automated vocal analysis reveals about the vocal production and language learning environment of young children with autism", *J. Autism Dev. Disord.,* vol. 40, no. 5, pp. 555-569, 2010.
[http://dx.doi.org/10.1007/s10803-009-0902-5] [PMID: 19936907]

[13]   N.I. Chongjia, L. Wenju, and X. Bo, "Research on large vocabulary continuous speech recognition system for mandarin Chinese", *J. Chin. Inform. Process.,* vol. 23, no. 1, pp. 112-123, 2009.
[http://dx.doi.org/10.3969/j.issn.1003-0077.2009.01.018]

[14]   D. Wang, and S.S. Narayanan, "Robust speech rate estimation for spontaneous speech", *IEEE Trans. Audio Speech Lang. Process.,* vol. 15, no. 8, pp. 2190-2201, 2007.
[http://dx.doi.org/10.1109/TASL.2007.905178] [PMID: 20428476]

[15]   W.U. Yanhua, "Analysis and application of short-time magnitude and short-time zero-crossing rate in speech", *Comp. Knowl. Technol.,* vol. 5, no. 33, 2009.
[http://dx.doi.org/10.3969/j.issn.1009-3044.2009.33.082]

[16]   M. Jalil, F.A. Butt, and A. Malik, "Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals", *Technol. Adv. Electric. Electron. Comp. Eng. (TAEECE),* pp. 208-212, 2013.

[17]   D. Xu, J. Gilkerson, J. Richards, U. Yapanel, and S. Gray, "Child vocalization composition as discriminant information for automatic autism detection", In: *31st Annual International Conference of the IEEE EMBS,* Minneapolis, 2009, pp. 2-6.

[18]   D. Lu, and P. Zhou, "Research on speech endpoint detection and initial/final segmentation based on the dual-threshold algorithm", *J. Guilin Univ. Electron. Technol.,* vol. 6, pp. 480-484, 2011.