SPEECH EVALUATION WITH SPECIAL FOCUS ON CHILDREN SUFFERING FROM APRAXIA OF SPEECH

Manasi Dixit[1] and Shaila Apte[2]

[1]Department of Electronics K.I.T.'s College of Engineering, Kolhapur,416234
Maharashtra, India
mrdixit@rediffmail.com
[2]Department of Electronics and Communication Engineering,
Rajarshi Shahu College of Engineering, Pune, Maharashtra,India
sdapte@redifmail.com

ABSTRACT

Speech disorders are very complicated in individuals suffering from Apraxia of Speech-AOS. In this paper , the pathological cases of speech disabled children affected with AOS are analyzed. The speech signal samples of children of age between three to eight years are considered for the present study. These speech signals are digitized and enhanced using the using the Speech Pause Index, Jitter,Skew ,Kurtosis analysis This analysis is conducted on speech data samples which are concerned with both place of articulation and manner of articulation. The speech disability of pathological subjects was estimated using results of above analysis.

KEY WORDS

Speech-Pause Index (SPI), Lexical Stress Index (LSI), Fundamental Frequency Analysis, Speech Intensity Analysis, Coefficient of Variation Ratio (CVR), Formant Analysis ,Jitter,Skew,Kurtosis ,Consonant production errors ,Speech intelligibility, Speech-recognition,

1. INTRODUCTION

Apraxia of Speech(AOS) is caused by a neurologic disorder or injury (e.g. cerebral palsy or traumatic brain injury).. In this present work, the speech data utterances by children of age group of 3 to 10 years were recorded and digitized. The digitized signal was further processed by using a MATLAB platform. The speech data was analyzed to check the disorders due to AOS. The acoustic analysis (e.g., spectrography,) is conducted on speech data. Perceptual assessment of speech is done so as to get important information regarding articulation, resonance and speech intelligibility[1]. Perceptual assessment provides important information regarding misarticulation, resonance and speech intelligibility, In this present work ,it is observed that the speech disabled children produce misarticulation errors mainly due to weak motor movements and inappropriate closure of oral cavity uncoordinated articulator movements, breathing problems[6,7]. When speech is affected, all aspects of speech production may be affected including respiration, phonation, resonation, articulation, and prosody due to Rigidity of motor muscles , Velopharyngeal incompetency, Auditory processing, and Language impairments[2,10]. The paper is organized as follows. Section 2 deals with the literature review .Section 3 discusses the methodology for experimental work and section 4 presents the results.

DOI : 10.5121/sipij.2013.4305

2. LITERATURE REVIEW

According to R.Cmelja, J. Rusz the changes of vocal cords vibration are perceived as changes of pitch period or fundamental frequency in voice. The fundamental frequency is influenced by properties of the vocal cords for example elasticity, weight and length. Thus the variations in fundamental frequency provide instability measurements and work as markers about the condition and functionality of the vocal cords. Abnormal characteristics of the fundamental frequency are found in a number of voice and speech disorders .Frequency instability or jitter is derived from the instantaneous values of pitch periods[1] . John-Paul Hosom, Lawrence Shriberg, Jordan R. Green suggested the method of estimation of Jitter . They have also discussed the impact of lexical stress and coefficient variation ratio in case of evaluation of pathological voices [2,10]. Darush Mehta discussed the liear source filter model of the vocal tract and the variations in formant frequencies act as a diagnostic marker in case of speech diabled subjects[4]. Skew and Kurtosis factors were evaluated using Praat.

3. METHODOLOGY

The present work is based on study of children with Marathi mother tongue. The speech data of normal subjects/children and pathological subjects/children of the same age group between 3 to 10years is collected. The children were trained to utter similar words before recording. The speech data consists of isolated words, connected words , fast uttered sentences and songs for eg. Prarthana-School-Prayer, National anthem and Pledge,Nursery Rhymes,famous film songs etc. The speech data was recorded using Sony Intelligent Portable Ocular Device (IPOD) in digital form. The recording was carried out in a pleasant atmosphere and maintaining the children in tension-stress free environment. The recorded signal is transformed into .wav file by using GOLDWAVE software. The isolated word speech data is given in Table-1. The data was collected at Chetana Vikas Mandir, a special school established to educate mentally retarded children as well as children with various disorders. It is located at Kolhapur, India.

3.1. Segmental Analysis

The Segmental Analysis was carried out for utterances of particular isolated words as well as complete sentences and paragraph /songs. The utterances made by 20 normal subjects and 10 pathological subjects were analyzed. Proposed diagnostic markers for suspected Apraxia of Speech (AOS) are vowel duration analysis, vowel formant analysis, and consonant transition index (CTI) analysis were carried out. Various misarticulation cases were analyzed in case of pathological subjects. The spectrograms were studied for formant analysis[1,4]. Fast uttered words or continuous sentences exhibit greater complexities with respect to speech intelligibility. The results obtained with the help of proposed MATLAB Algorithm were verified using Praat and SFS Open source softwares.

3.2. Suprasegmental Analysis

The Suprasegmental Analysis was carried out for utterances of particular isolated words as well as complete sentences and paragraph /songs. The utterances made by 20 normal subjects and 10 pathological subjects were analyzed. Proposed diagnostic markers for suspected Apraxia of Speech (AOS) are Speech-Pause Index (SPI), Fundamental Frequency Analysis[9], Speech Intensity Analysis ,Jitter[11],Skew and Kurtosis analysis[1,2,10].

Skew and Kurtosis analysis represents the deviations due to stressed utterences compared to the unstressed utterences in normal subjects as obtained from a speaker's productions for isolated

word and sentences forms. Composite weightings for the three stress parameters were determined from a principal components analysis. The Coefficient of Variation Ratio (CVR) expresses the average normalized variability of durations of pause and speech events that were obtained from a conversational speech sample. These diagnostic markers address temporal variations observed in the speech of children with suspected AOS [2,10].

3.3 Suprasegmental factors concerned with Marathi language

The present study uses marathi words and sentences. The following factors are of concern for marathi language.

Length: Vocalic length is mostly predictable. With the exception of ə, the last vowel of a word is long unless the vowels are followed by a combination of consonants such as nt, tr, kt. The length is phonemic in i, u.

Nasal vowels: Use of nasal vowels as independent entities and they vary from speaker to speaker. They are found in certain adverbs, nouns, and plural nouns in the context of case and postpositions. They are phonemic in certain dialects.

Accent: Marathi is said to have a stress accent. Length, pitch, and sonority play a role in determining the loudest accent.

Phonology :Modern Marathi script, known as' balbodhi', is based on the Sanskrit Devnagari script, with certain modifications. Unlike English, Devnagari is alpha syllabic. It uses certain diacritics for vowels when combined with consonants. The diacritics distinguish long and short vowels. There are special systems to denote consonant clusters.

Traditional Marathi alphabetic chart lists 16 vowels and 36 consonants based on Sanskrit. Today, many of these alphabets are obsolete. Modern Marathi has 8 basic vowels and 34 consonants including two semivowels. Table 1 and Table 2 indicate the vocalic and consonantal charts and their respective features.

Table 1. Devanāgarī alphabet for Marathi Vowels and vowel diacritics with IPA codes

अ	आ	इ	ई	उ	ऊ	ए	ऐ	ओ	औ	अं	अः	ऋ	ऌ
a	a	ɪ	ɪ	u	u	e	aɪ	o	au	aɲ	aɳ	r̥	l̥
[ə]	[a, ai]	[i]	[i:]	[u]	[u:]	[e]	[ai]	[o]	[au]	[aŋ]	[əh]	[ɾ̩]	[l̩]
प	पा	पि	पी	पु	पू	पे	पै	पो	पौ	पं	पः	पृ	पॄ
pə	pā	pi	pī	pu	pū	pe	pai	po	pau	paṅ	paḥ	pr̥	pl̥

Table 2. Devanāgarī alphabet for Marathi Consonants with IPA codes

The isolated word speech data emphasized in the present work is described below. It indicates case studies of different types of Articulation errors with respect to both place of articulation and manner of articulation[6,7]. Table 3 lists the place of articulation and manner of articulation for marathi words. Table 4 lists the fundamental frequencies for normal and pathological subjects and Table 5 lists the second and third formants for normal and pathological subjects. Table-6. Lists the Speech Intensity,Jitter,Speech –Pause Analysis for Different Subjects.

Table-3. List of specific place of articulation and manner of articulation of Consonants

Sr. No.	Place of articulation	Manner of articulation	English Letters	Marathi Letters	Marathi Words
1	bilabial	plosive	p,b	प फ ब भ	आपण फजिती बरी भाव
	bilabial	nasal	m	म	मामा
	bilabial	approximant	w	व	वासू
2	labio-dental	fricative	f, v	फ व्ह	फुरसे व्हय
3	dental	fricative	th, <u>th</u>	त थ द ध	तारा थोडी दूध
4	alveolar	plosive	t, d	ट ठ ड ढ	डॉक्टर ढग
	alveolar	nasal	n	न ण	बाण कान
	alveolar	fricative	s, z	स	साप
	alveolar	approximant	r/l	र ल ळ	रसाळ लळा
5	post-alveolar	fricative	sh,zh	ष श झ	षटकोन शहामृग झगा
	post-alveolar	affricate	ch,j	च छ ज	चंदा छुमछुम जहाज
	post-alveolar	approximant	y	य	योयो
6	velar	plosive	kg	क ख ग घ	काय ग ख्रातेस घर
	velar	nasal	ng	ङ.	वाड.मय
7	glottal	fricative	h	ह	हवा
Other Isolated Words Used for Analysis			गह्य अनंत संयुक्त राष्ट अज्ञानी हृदय		

Table-4.Fundamental Frequency Analysis - f_0 and variations of f_0 for different subjects

Sr.No.	Subject	f_0	Variations of f_0	Remarks
1	Speaker 1	255 Hz	96.7-319	Normal Female
2	Speaker 2	228.7Hz	141.3-319.6	Normal Female
3	Speaker 3	210.55 Hz	60-307.7	Patological Female Subject
4	Speaker 4	188.8 Hz	129.7-309.6	Normal Male
5	Speaker 5	198.8 Hz	129-265.6	Pathological Male Subject
6	Speaker 6	139.75 Hz	101-297.9	Pathologic Male Subject
7	Speaker 7	181.5 7Hz	77-297.9	Pathologic Male Subject
8	Speaker 8	159.54 Hz	82-315.5	Pathologic Male Subject
9	Speaker 9	163.4 Hz	69-319	Pathologic Male Subject
10	Speaker 10	240.21 Hz	91-319	Pathologic Male Subject

Table-5. Formant Analysis- f_1, f_2 and f_3 for Different Subjects

Sr.No.	Subject	f 1 -Hz	Bw of f 1	f 2 -Hz	Bw of f 2	f 3 -Hz	Bw of f 3
1	Speaker 1	580.94	379.6	1994.12	1349.86	2911.28	805.6
2	Speaker 2	618.21	154.03	1733.73	100.85	2740.9	351.59
3	Speaker 3	487.95	194.31	1530.16	143.32	2450.86	157.39
4	Speaker 4	640.57	70.6	1760.14	107.98	2738.32	261.8
5	Speaker 5	563.04	703.13	1397.31	372.92	2828.74	355.64
6	Speaker 6	811.47	503.09	1632.88	1496.1	2683.95	1069.35
7	Speaker 7	700.87	274.05	1598.82	1186.99	2548.72	749.50
8	Speaker 8	659.90	319	1751.96	424.39	2908.14	314.45
9	Speaker 9	1055.46	46.11	1875.62	202.5	2746.21	400.12
10	Speaker 10	1026.46	29.18	1761.64	1075.86	2561.33	247.57

Table-6. Speech Intensity,Jitter,Speech –Pause Analysis for Different Subjects

Sr. No.	Subject	Sound Amplitude mean-pascal	Sound Mean Power -dB	Total Energy in Air-Joules/ m2	Intensity Variation-dB	Jitter %	Speech-Pause Index %	Skew	Kurtosis
1	Speaker 1	0.41574	84.82	0.00660	60.72-89.53	17.16	38.07	4.32	22.70
2	Speaker 2	0.05110	84.34	0.0063	55.9-91	27.28	27.06	4.2	16.81
3	Speaker 3	-0.00397	70.02	0.00051	45-79.98	13.02	64.55	2.47	4.76
4	Speaker 4	-0.0523	72.63	0.00028	56.87-84.66	27.38	40	3.4	12.86
5	Speaker 5	0.30647	85.12	0.04907	60.38-91.74	31.4	62.93	2.24	5.28
6	Speaker 6	0.38278	87.91	0.01469	69.38-90.59	22.77	25.63	3.42	11.10
7	Speaker 7	0.3592	83.57	0.00760	56.28-87	2.29	48.8	2.79	6.47
8	Speaker 8	0.0001	88.41	0.16007	60.41-91.50	35.8	21.85	2.99	9.41
9	Speaker 9	0.00018	83.67	0.0166	58.45-84.4	28.06	42.22	1.97	2.66
10	Speaker 10	0.00018	83.67	0.0166	38.54-86.56	33.45	54.55	1.75	4.09

4. CONCLUSION

The pathological subjects affected with Apraxia of speech exhibit rigidity in motor movements. Various types of misarticulation errors occur in different subjects. The observations are weak consonant production and inappropriate closure of oral cavity while producing Bilabial plosives. Utterances of some of the plosives and alveolar approximants were not possible in case of some of the subjects. The Segmental and Suprasegmental Analysis indicators Speech-Pause Index (SPI), was found to be high in pathological subjects. The Skew factor is observed below the threshold level of 3.4 and the Kurtosis factor is observed below the threshold level of 12 in almost all pathological subjects. High Pitch values falling in the female pitch range are observed in case of pathological male subjects. The Formants are widely spread in case of pathological subjects. High range of itensity variations,high % Jitter levels, high values of speech-pause index(SPI),low skew levels and low level of Kurtosis factor is the diagnostic marker set of pathological male/female speech.

ACKNOWLEDGEMENTS

The authors would like to thank the participants of research work- all the mentally retarded children with speech disorders and the Director of the special school Chetana Vikas Mandir Mr.Pawan Khebudkar.

REFERENCES

[1] R.Cmelja, J. Rusz "Laboratory Tasks From Voice Analysis in the Study of Biomedical Engineering using MATLAB" Research Program MSM6840770012 Trans Disciplinary Research in Biomedical Engineering-Voice and Speech Disorders, Charles university ,Prague, Czech Technical university ,Prague , Dallas university

[2] John-Paul Hosom, Lawrence Shriberg, Jordan R. Green"Diagnostic Assessment of Childhood Apraxia of Speech Using Automatic Speech Recognition (ASR) Methods" J Med Speech Lang Pathol. 2004 December ; 12(4): 167–171.

[3] Marcelo de Oliveira Rosa, José Carlos Pereira, and Marcos Grellet" Adaptive Estimation of Residue Signal for Voice Pathology Diagnosis" IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 47, NO. 1, JANUARY 2000

[4] Darush Mehta "Aspiration Noise During Phonation: Synthesis, Analysis, And Pitch -Scale Modification" MASSACHUSETTS INSTITUTE OF TECHNOLOGY February 2006

[5] Lingyun Gu, John G. Harris,et.al. "Disordered Speech Assessment Using Automatic Methods Based on Quantitative Measures " EURASIP Journal on Applied Signal Processing 2005:9, 1400–1409 c_ 2005 Hindawi Publishing Corporation

[6] Oscar Saz,1 Javier Sim´ on,2 W.-Ricardo Rodr´ıguez,1 Eduardo Lleida,1 and Carlos Vaquero1 *Research Article*"Analysis of Acoustic Features in Speakers with Cognitive Disorders and Speech Impairments" Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2009, Article ID 159234, 11 pages doi:10.1155/2009/159234

[7] Juan Ignacio Godino-Llorente,1 Pedro G´omez-Vilda (EURASIPMember),2 and Tan Lee3 *Editorial* "Analysis and Signal Processing of Oesophageal and Pathological Voices" Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2009, Article ID 283504, 4 pages doi:10.1155/2009/283504

[8] CatherineMiddag,1 Jean-PierreMartens,1 Gwen Van Nuffelen,2 andMarc De Bodt2 *Research Article* "Automated Intelligibility Assessment of Pathological Speech Using Phonological Features "Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2009, Article ID 629030, 9 pages *doi:10.1155/2009/629030*

[9] Arundhati S. Mehendale and Manasi R. Dixit "Speaker Identification" Signal & Image Processing : An International Journal (SIPIJ) Vol.2, No.2, June 2011, DOI : 10.5121/sipij.2011.2206

[10] Lawrence D. Shriberg "Studies of Childhood Apraxia of Speech in Complex Neurodevelopmental Disorders" Conference on Motor Speech: Motor Speech Disorders ,Monterey, California,March 6-9, 2008

[11] D´arcio G. Silva, Lu´ıs C. Oliveira, andM´ario Andrea "Jitter Estimation Algorithms for Detection of Pathological Voices" Research Article Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2009, Article ID 567875, 9 pages doi:10.1155/2009/567875

Authors

Mrs.Manasi Dixit is working as Associate Professor in KIT's College of Engineering,Kolhapur .Her teaching experience is 28 years. Her main fields of interest are Digital Signal Processing,Speech Processing,Image Processing and Microwave Engineering. 16 PG students have completed their research work and have been awarded M.E.(E&TC) degree under her guidance. She has published 14 papers in reputed international journals.She has worked in the capacity of the SENATE member Shivaji University,Kolhapur and BOS-Board of Studies Member for Electronics and Telecommunication Engineering, Shivaji University,Kolhapur.

Prof. Dr. Shaila Dinkar Apte is currently working as a professor on PG side in Rajarshi Shahu College of Engineering, Pune and as reviewer for the International Journal of Speech Technology by Springer Publication, International Journal of Digital Signal Processing, Elsevier Publication. She is currently guiding 5 Ph.D. candidates. Four candidates have completed their Ph.D. under her guidance. About 50 candidates completed their M.E. dissertations under her guidance. Almost all dissertations are in the area of signal processing. She has a vast teaching experience of 32 years in electronics engineering, and enjoys great popularity amongst students. She has been teaching Digital Signal Processing and Advanced Digital Signal Processing since last 18 years. Her previous designations include being an Assistant Professor in Walchand College of Engineering, Sangli, for 27 years; a member of board of studies for Shivaji University and a principle investigator for a research project sponsored by ARDE, New Delhi. She has published 28 papers in reputed international journals, more than 40 papers in international conferences and about 15 papers in national conferences. She has a patent published to her credit related to generation of mother wavelet from speech signal. Her books titled "Digital Signal Processing" and "Speech and Audio Processing" are published by Wiley India. A second reprint of second edition of the first book is in the market. The book titled "Advanced Digital Signal Procesing" will be published recently in June 2013 by Wiley publishers.