31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Assessment of Laryngeal Dysfunctions of Dysarthric Speakers

V. Surabhi, P. Vijayalakshmi, Steffina Lily. T and Ra. V. Jayanthan

Abstract—**Dysarthria is a neuromotor impairment of speech that affects one or more of the speech sub-systems. It is reflected in the acoustic characteristics of the phonemes as deviations from their healthy counterparts. In the current work, the deviations associated with laryngeal dysfunctions are analysed in order to assess and quantify parameters that will help evaluate dysarthria. Perturbation measures, pitch period statistics and Pitch Variation Index (PVI) are computed for the assessment of laryngeal dysfunctions of dysarthric speakers. The assessments were performed on the Nemours database of dysarthric speech and compared with normal speakers available in the TIMIT speech corpus. The results were correlated with Frenchay Dysarthria Assessment (FDA) scores. The analysis resulted in a technique to predict the degree of severity of dysarthria and illustrate the multi-causal nature of the disorder.**

I. Introduction

DYSARTHRIAS are complex neuromotor impairments of speech that can affect every component of speech production. In most individuals, the disruption is spread over more than one sub-system; the respiratory, laryngeal and articulatory. This is decided by which of the dimensions of speech such as voice quality, intelligibility and prosody are affected, thus hindering effective communication. Hence it is mandatory to assess speech subsystems to enhance communication of dysarthric speakers [1].

Acoustic analysis can be informative since it provides quantitative analyses that carry potential for sub-system description and holds key details regarding speech rate, articulatory configuration for vowels and consonants, rates of change in overall configuration of vocal tract, flexibility of articulatory behaviour and aspects of phonatory behaviour. It helps determine the correlates of perceptual judgements of intelligibility, quality and type of dysarthria. Therefore, acoustic analyses can be a valuable complement to perceptual judgement. Although perceptual judgement has been the primary means for classification and description, the reliability and validity are a question since these are performed by different specialists with no common training in perceptual rating [2]. Also, perceptual judgements alone cannot discriminate between disruptions that occur simultaneously in two or more speech sub-systems.

Kent, Weismer and Rosenbek designed a profile to direct research on acoustic or physiologic correlates of dysarthric intelligibility impairment in addition to development of an intelligibility test for dysarthria [3]. Hess proposed the fundamental frequency over existing time and frequency domain parameters for acoustic analysis of dysarthria [4]. Weismer, Jeng, Laures, Kent and Kent analyzed acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders and deduced that temporal variables could differentiate some disorders while not others. Also, that the relation between acoustic measures and scaled speech intelligibility was complex, and could be expressed in terms of sentences vs. single-word intelligibility estimates and their underlying acoustic bases [5]. Herzel, Berry, Titze and Saleh applied fractals and chaos theory for acoustic analysis of voice and its disorders [6]. However, the progress in acoustic studies of dysarthria has been slow owing to several factors, including: (1) the relatively modest research effort given to neurogenic speech disorders, (2) the difficulty of acoustic analysis for speakers who may have phonatory disruptions, hypernasality, imprecise articulation and other properties that confound acoustic description, and (3) rather few published examples of broadly directed acoustic analyses of dysarthria [2].

In this work, phonemes were extracted from the Nemours and the TIMIT speech corpora for the analysis. The speech rate was calculated for all the phonemes. Then acoustic analysis was performed on the vowels (voiced sounds). This involved estimation of acoustic parameters afore-mentioned and their comparison with the normal speakers. The latter consisted of normal speakers from the TIMIT speech corpus. Statistical, quantitative and graphical analyses were evaluated. The reliability and validity of the underlying inferences projected were estimated by corroboration with FDA. Correlation was calculated to quantitatively measure the success rate of the assessment.

The paper is organized as follows. The speech corpora used have been explained primarily, followed by elucidation of the methodology adopted in this work. Correlation with FDA has then been presented. The last section contains inferences and decisively, the success rate of the analysis.

II. Speech Corpora

A. TIMIT Speech Corpus

For this work, the speech data from TIMIT [7] speech corpus is considered as the reference data. It consists of hand-labelled and segmented data of quasi-phonetically balanced sentences read by 630 native speakers of American English. The data used for the acoustic analysis are the phonemes: /a/, /i/ and /u/ from the New England dialect

Manuscript received June 10, 2009.

V. Surabhi, T. Steffina Lily and Ra. V. Jayanthan are now undergraduate students of Biomedical Engineering at SSN College of Engineering, Chennai, India – 603110 (e-mail: pkv.surabhi@yahoo.com, steffinalily.t@gmail.com, vigneshr_bio@yahoo.co.in).

Dr. P. Vijayalakshmi is now a Professor, Department of Electronics and Communication Engineering at SSN College of Engineering, Chennai, India (e-mail: vijayalakshmip@ssn.edu.in).

while the test and training data were utilized for speech rate analysis.

B. Nemours Speech Corpus

The database used for the assessment of dysarthria consists of 10 dysarthric speakers' and one normal speaker's speech data from the Nemours database of dysarthric speech [8]. With this corpus time–aligned phonetic transcriptions are given for all the dysarthric speakers' speech data. This database contains Frenchay dysarthria assessment (FDA) [9] scores for 9 dysarthric speakers. FDA is a well-established test for the diagnosis of dysarthria. The test is divided into 11 sections, namely, reflex, palate, lips, jaw, tongue, intelligibility etc. Each dysarthric speaker is rated on a number of simple tasks. FDA is a 9 point scale with scores ranging from 0 to 8. In FDA, a score of '8' represents normal function and '0' represents no function.

The dysarthric speakers considered for the present work are affected mostly with spastic dysarthria. The characteristics of spastic dysarthria are: (a) The voice of a patient with spastic dysarthria is sometimes strained or strangled. (b) Pitch is low, with pitch breaks occurring in some cases. (c) Hypernasality typically occurs. (d) Range of movement, tongue strength, speech rate and voice onset time for stops are reduced. (e) There is an increase in phoneme to phoneme transitions, in syllable and word duration, and in voicing of unvoiced stops. Of the 10 speakers, speaker FB was found to be severely dysarthric to an extent that segmentation was not feasible. Hence FB could not be subjected to acoustic analysis.

III. Methodology

The assessment was carried out as follows:

A. Durational and Speech Rate Analysis

From the time-aligned phonetic transcription available with the Nemours and the TIMIT database, the duration of each phoneme for each of the speakers is computed. The means and variances of the duration of dysarthric speech are normalized with respect to the normal speech. From these calculations, it is observed that the normalized mean duration of dysarthric speakers over all the phonemes is always greater than (≈twice) that of the normal speakers. Apart from duration, the speech rate (number of phonemes uttered per second) for each dysarthric speaker and the normal speakers (TIMIT) are calculated. The speech rate of each dysarthric speaker is normalized with respect to the mean speech rate of normal speakers. From this, it is observed that the speech rate of dysarthric speakers varies from a minimum of 1.5 times to a maximum of 2.8 times less than that of normal speech. This may be reflected in the variation in the pitch period and hence an acoustic analysis is conducted to verify this [10].

B. Acoustic Analysis

Segmented speech segments of voiced sounds were extracted from the corpora. A range of acoustic variables were extracted. These included pitch period, variance in pitch period, pitch variation index, jitter and shimmer. The speaker information content procured were classifiable as:

1) Frequency Parameters: Frequency measures give information regarding dysarthric subject's habitual pitch, optimal pitch and pitch range, the degree of pitch steadiness and any pitch alterations during speech. Alteration in laryngeal function resulting from changes in vocal fold elasticity, stiffness, length or mass can affect the fundamental frequency. Consequently, statistical pitch period (mean and variance) and Pitch Variation Index (PVI) were useful for identifying presence of abnormal laryngeal function and monitoring laryngeal function over time. These were indicators of identification of the presence of either reduced or excessively variable pitch use.

2) Perturbation Parameters: Acoustic perturbation measures reflect short-term and long-term variability in the fundamental frequency, Fo, and amplitude. Specifically, measurement of short-term, cycle-to-cycle variability in Fo is 'jitter', while short-term variability in amplitude is 'shimmer'. Two further measures each were evaluated for all the dysarthric speakers available in the database to enhance speaker information detail. These were: Jitter (local), Jitter (Relative Average Perturbation/ RAP), Shimmer (local) and Shimmer (DDP). Both jitter and shimmer reflected high-frequency fluctuations; consequently, alterations in these values may be strongly associated with laryngeal tissue abnormalities, asymmetries in movement and fast-acting neuromuscular fluctuations. Specifically, the presence of jitter indicated variations in vibratory patterns of the vocal chords while shimmer indicated vocal fold instability. With respect to perceptual features, abnormalities in jitter and shimmer were closely associated with judgements of voice hoarseness, roughness and harshness.

IV. Observations

The acoustic variables were quantified and compared with FDA. Pearson's correlation coefficient was calculated to establish the degree of corroboration between this work and FDA.

A. Speech Rate Analysis

The speech rate analysis proved to match with perceptual judgement to a superlative degree as is seen in Fig. 1. While both exhibit the same pattern for the majority, speakers RL and SC alone show exception. Normal speakers converse at an average of 15 phonemes/second.

B. Acoustic Analysis

1) General Patterns: Certain common patterns in parameter variation were exhibited. Primarily, speech rate varied linearly with FDA laryngeal time rating. Similarly, PVI exhibited linear change with FDA laryngeal pitch rating while pitch period was inversely related.

2) Subject–Specific Analysis: However, it was obvious that no standard patterns could be evolved to correlate acoustic parameters, speech rate and perceptual judgement (FDA ratings). The fact that each parameter showed individualistic variations, when corroborated with FDA,

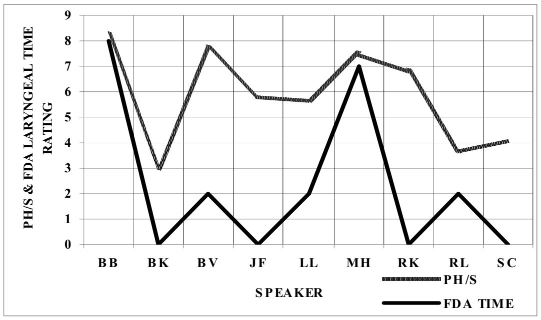


Fig. 1. Comparison of phonemes/second and FDA Laryngeal Time rating for every dysarthric speaker

reinforced the fact the dysarthria may be caused by multiple sub-system disorders in specific combinations. Hence, a speaker specific analysis was necessitated. Analyzing Fig. 1, Fig, 2, Fig. 3 and Fig. 4:

BB: Frequency parameters including pitch period, variance and PVI were of moderate nature, his pitch range going upto 7ms. His jitter was mostly within the pathological threshold while his shimmer clustered around 5-10% indicating voice instability. He had highest speech rate of 8ph/s. Correspondingly, laryngeal time and speech ratings by FDA were 8 with good overall assessment and word intelligibility 4, thus correlating with this work.

BK: BK's frequency parameters were in the normal range. The pitch period ranged from 1ms to 10ms (i.e. 9ms). Variance in pitch period was extremely low, reflecting in the probability of vocal fold stiffness causing excess pitch steadiness and lack of modulation. Correspondingly, FDA laryngeal pitch rating was 0. His perturbation parameters

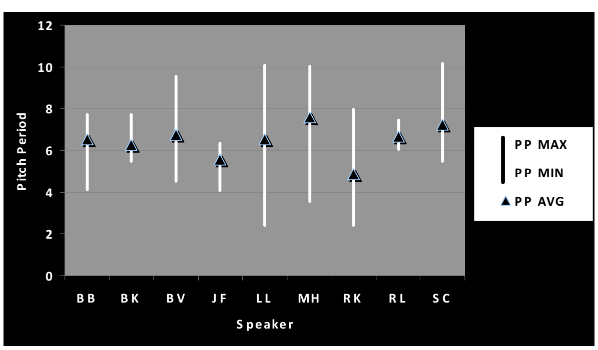
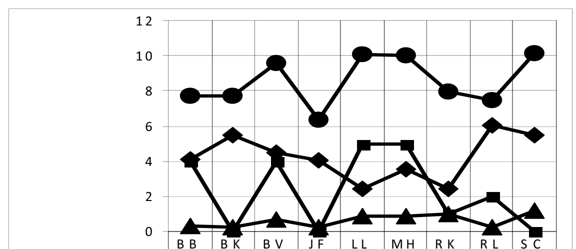


Fig. 2. Maximum and minimum pitch periods with average pitch period for every dysarthric subject for voiced phoneme /aa/. ('PP' denotes Pitch Period)

were consistent. Jitter values varied around 1% while the shimmer went upto 11%, thus reflecting consistent laryngeal tissue abnormalities, symmetric movement and fast-acting neuromuscular fluctuations. Vocal folds could hence be relatively stable. However, BK was the slowest of the 10 dysarthric speakers available in the database. His speech rate was a meagre 3 ph/s being totally incomprehensible. This correlated with FDA intelligibility rating of 0. FDA overall ratings were all extremely slow. However, his jaw and palate were moderate in speech. Also, FDA summary states that: Produced /p/ with labiodental contact that can lead to slight build-up of pressure, lots of compensatory articulation. It is on these premises that perceptual rating for BK is low. Hence, it can be concluded that jaw and palate being the important components of vowel production, our results agree with FDA.

BV: BV's frequency parameters showed a moderately high level of variability in terms of pitch period, PVI and variance. Equivalently, FDA laryngeal pitch rating was 4. He had 2nd highest speech rate of 8 ph/s. However, FDA time rating was quite low at 2. He showed considerable variation in jitter within the range 1-5% and shimmer reaching 35%. This could be the effect of vocal fold instability. Correspondingly, FDA laryngeal and respiratory speech ratings were at 2 and his word intelligibility was 0.

JF: JF exhibited a moderately stable pitch period, variance and PVI with a moderate range (6ms). He had a moderate speech rate of 6ph/s. His perturbation parameters were also considerably stable, jitter averaging around 1% while shimmer around 5%. FDA assessment correlates by its



	BB	BK	BV	JF	LL	MH	RK	RL	SC
PP MAX	7.69	7.69	9.55	6.33	10.1	10	7.97	7.43	10.1
PP MIN	4.11	5.46	4.49	4.07	2.41	3.55	2.41	6.03	5.46
PVI	0.30	0.26	0.68	0.24	0.86	0.84	0.99	0.24	1.17
FDA PITCH	4	0	4	0	5	5	1	2	0

Fig. 3. Comparison of Pitch Period (Maximum), Pitch Period (Minimum), Pitch Variation Index and FDA Laryngeal Pitch Rating for individual dysarthric subjects for voiced phoneme /aa/.

judgement that respiratory speech, jaw, palate and tongue are moderate while laryngeal function is severely affected. Moderate intelligibility of 4 prevailed. Hence, JF is a moderate speaker, especially in terms of vowels.

LL: LL possessed widest pitch period range of 13ms (1 to 14ms). Consolidating, his pitch period and PVI are moderate. His perturbation parameters were also considerably stable, jitter around 1-3% while shimmer averaging around 5%. His probable periodic vocal fold instability reflected in sudden jerks in variance and range. He had a moderate speech rate of 6ph/s. FDA laryngeal time was at a considerable low of 2. All his FDA ratings were moderately high with intelligibility of 4.

MH: Frequency parameters of MH showed large variations. His pitch period, range (10ms), PVI and variance were high. Jitter averaged around 1% while shimmer around 5-10% indicating high vocal fold instability. He had a moderate speech rate of 7.5ph/s. On the contrary, his perceptual ratings were very good. But FDA comment that

MH has difficulties with vowel distinctions justifies the assessment in this work.

RK: Frequency parameters shown by RK exhibited a high pitch period range of 10µs, moderate variations in pitch period reaching 5 in 1 instance with considerable variance. His jitter lay between 1-10% and shimmer averaged 5-15%. RK's speech rate was a considerable 7ph/s. FDA assessed that respiratory speech, jaw and palate were moderate while laryngeal time was 0, pitch and speech 1. He was moderately intelligible with an FDA score of 4. FDA summary states that RK's intelligibility evaluation is based only on the words and sentences he attempted to read. Hence, RK may possess problem only with unvoiced sounds, thus validating the analysis in this work.

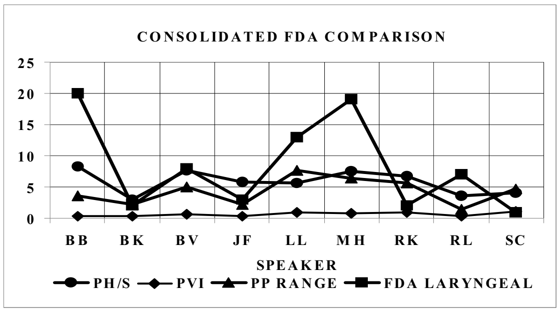


Fig. 4: Consolidated analysis of speech rate and acoustic parameters (PVI and Pitch Period Range) against consolidated FDA Laryngeal rating over 24 points (time, speech, pitch) for phoneme /aa/.

RL: RL's frequency and perturbation parameters were moderately good. His pitch period range was the lowest (6ms) indicating possible voice stiffness, roughness and hoarseness. He was found to be an extremely slow speaker at 4ph/s. However, FDA gives a moderately low assessment with a time rating of 2 and moderate intelligibility of 4.

SC: While SC was a very slow speaker at 4ph/s, his frequency parameters (PVI, variance and pitch period range - 14µs) exhibited a lot of variations. Jitter was centered around 1% while his shimmer indicated chronic voice instability and laryngeal tissue abnormalities ranging upto 30%. This correlated with FDA assessment that indicated chronic dysarthria with intelligibility 1. Vowels may be of special difficulty to him, confirmed by FDA summary stating that lot of tasks confounded by lack of air support. SC compensates for many deficiencies using alternate means of production.

C. *Correlation*

Pearson's correlation coefficient was deduced between FDA laryngeal assessment and the proposed assessment method. The results are as shown in Table 1.

V. CONCLUSION

Thus speech rate and acoustic analysis performed on dysarthric data predicted the degree of severity and deduced the causes. The success rate was established by reasonably high correlation coefficients obtained by comparing against the standard perceptual judgement modality, FDA.

TABLE I

CORRELATION BETWEEN FDA PARAMETERS AND ACOUSTIC PARAMETRS

FDA Parameter	Assessment Parameter	Correlation Coefficient
Laryngeal Time	Speech Rate	0.6425
Laryngeal (Total)	Speech Rate	0.6446
Laryngeal Pitch	PP Range	0.6184
Laryngeal Pitch	PP Maximum	0.5187
Laryngeal Pitch	PP Average	0.4605

PP = Pitch Period, Total = Pitch + Time + In Speech

REFERENCES

[1] P. Vijayalakshmi and M. R. Reddy, "Assessment of dysarthric speech and an analysis on velopharyngeal incompetence", *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York City, USA, Aug 30–Sept 3, 2006.

[2] Ray D. Kent, Gary Weismer, Jane F. Kent, Houri K. Vorperian and Joseph R. Duffy, "Acoustic studies of dysarthric speech: Methods, progress and potential", *Journal of Communication Disorders*, Vol. 32, Issue 3, pp. 141–186, 6 May 1999.

[3] Ray D. Kent, Gary Weismer, Jane F. Kent and Rosenbek, "Toward phonetic intelligibility testing based on perceptual analysis", *Journal of Speech and Hearing Disorders*, Vol. 54, pp. 482–499, Nov 1989.

[4] W. J. Hess, "Fundamental frequency determination of speech and singing voice – A review", *The Journal of the Acoustical Society of America*, Vol. 92, Issue 4, pp. 2428–2448, October 1992.

[5] H. Herzel, D. Berry, I. R. Titze and M. Saleh, "Analysis of vocal disorders with methods from nonlinear dynamics", *Journal of Speech and Hearing Research*, Vol. 37, pp. 1008–1019, October 1994.

[6] Gary Weismer, Jing-Yi Jeng, Jacqueline S. Laures, Ray D. Kent and Jane F. Kent., "Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders", *International Journal of Phoniatrics, Speech Therapy and Communication Pathology*, 53(1), 2001.

[7] W. M. Fisher, G. R. Doddington and K. M. Marshall, "The DARPA speech recognition research database: specifications and status", in *proc. DARPA workshop on Speech Recognition*, pp. 93–99, February 1986.

[8] X. M. Pedal, J. M. Polikoff, S. M. Peters, J. E. Leonzio, and H. T. Bunnell, "the Nemours database of dysarthric speech", in *Proceedings of Int. Conf. Spoken Language Processing,* Philadelphia, PA, USA, pp. 1962 – 1965, 1996.

[9] P. Enderby, "Frenchay dysarthric assessment", *British Journal of Disorders and Communications*, vol. 15, no. 3, pp.165-173, 1980.

[10] P. Vijayalakshmi, T. Nagarajan and M. R. Reddy, "Assessment of articulatory and velopharyngeal sub-systems of dysarthric speech", *Biomedical Soft Computing and Human Sciences*,(Accepted, due for publication in Jun 2009.)