Methods for Detection of Speech Impairment Using Mobile Devices

Taha Khan and Jerker Westin*

Academy of Industry and Society, Computer Engineering, Dalarna University, Borlänge, Sweden

Received: 15 February 2011; Revised: 22 March, 2011; Accepted: 10 April, 2011

Abstract: Speech impairment is an important symptom of Parkinson's disease (PD). This paper presents a detailed systematic literature review on speech impairment assessment through mobile devices. A two-tier review methodology is utilized. The first tier focuses on real-time problems in speech detection. In the second tier, acoustics features that respond to medication changes in Levodopa responsive PD patients are investigated for recognition of speech symptoms. The investigation of the patents reveals that speech disorder assessment can be made by a comparative analysis between pathological acoustic patterns and the normal acoustic patterns saved in a database. The review depicts that vowel and consonant formants are the most relevant acoustic parameters to reflect PD medication changes. Since consonants have high zero-crossing rate (ZCR) whereas vowels have low ZCR, enhancements in voice segmentation can be done by inducing ZCR. Our synthesis further suggests that wavelet transforms have potential for being useful in real-time voice analysis for detection and quantification of symptoms at home.

Keywords: Parkinson's disease, hypokinetic dysarthria, voice recognition, speech impairment, telemedicine.

1. INTRODUCTION

Since Parkinson's disease (PD) is progressive, there is a need of following up treatments over time. Medicine dosing needs to be adjusted daily with respect to physical exercise, food intake and mood [1]. Only 3-4% of the patients receive timely treatment [2]. Reasons for this are that individuals may have physical limitations which make it difficult for them to come for treatment or they may not have easy access to treatment provided by a therapist.

High-performance modern mobile devices (MMD) with advanced sound processing capabilities has potential to tackle the challenge of accessibility. Clinician monitoring and feedback can be preserved using the MMD since it has the portability to adapt clinician directed treatment to home self training. Portable feedback devices for pathological individuals have been previously investigated. Zicker *et al*. [3] found that an individual with PD was able to modify her behavior when she received feedback from a device that her speech was too soft. Multisite treatment and data to evaluate the delivery of consistent treatment among clinicians can be acquired using a MMD [3].

1.1. PD Effects on Human Voice

PD is characterized by the loss of dopaminergic neurons in brain. This loss results in dysfunction of basal ganglia pathway which is an essential part of the circuitry that mediates motor and cognitive functions. As a result of dopamine loss in basal ganglia, there can be a number of motor symptoms such as rigidity, akinesia, bradykinesia, rest tremor, postural abnormalities, and speech dysfunction [4].

*Address correspondence to this author at the Academy of Industry and Society, Computer Engineering, Dalarna University, SE- 781 88 Borlänge, Sweden; Tel: +46 23 77 88 27; Fax: +46 23 77 80 50; E-mail: jwe@du.se

Physical symptoms that can occur in the limbs can also occur in the speech system. These symptoms are classified as hypokinetic dysarthria (HKD) [5]. "Dysarthria" refers to a speech disorder due to a change in muscle control. Hypokinetic means reduced movement. Thus, hypokinetic dysarthria is reduced movement of the muscles used for speech production.

1.2. Characteristics of Hypokinetic Dysarthria

HKD can affect respiration, phonation, resonation and articulation in speech [4-6]. Respiration problems disturb voice loudness in PD patients [6]. The reason is control of inhalation and exhalation enables a person to maintain adequate loudness of speech through a conversation. Persons with PD may speak on the "bottom" of his or her breath i.e. inhale, exhale, then speak; rather than on the "top" i.e. inhale, speak, exhale remaining air. PD individual's voice is an average of 2-4 dB softer than the normal voice [6].

Vocal folds vibration during phonation creates pitch of the voice. Vocal folds vibrate quickly during high-pitched sounds and vibrate slowly during low-pitched sounds [6]. Many individuals with PD notice changes in pitch of their voices [7]. Monotone or lack of vocal inflection or melody in voice is also a common complaint [8].

The resonating system determines richness of the voice. Soft palate, located in the back of the mouth roof closes off the nasal cavity while speaking, except when producing nasal sounds such as "...ing", "m..." or "n...". The soft palate does not move normally in PD. A nasal quality in voice is produced as the air is leaked into the nose due to the soft palate's inadequate movement [8].

The articulator system comprises of the face muscles, lips, tongue, and jaw. Imprecise articulation in PD is attributed to the reduced movement or lack of coordination of face muscles [7]. Imagine that your face was very cold

1877-6124/11 $100.00+.00

and it was difficult to move your facial muscles, your speech becomes slurred and unclear.

1.3. Research Challenges

Voice characterization in real-life environment is challenging. Speech datasets could possibly be collected *via* phone calls by the patients so that speech assessment can be more timely accurate and productive. But due to the background noise or in the medium of transmission, human voice detection becomes difficult. Distance of patient's mouth from the phone's mouth-piece may also create problem in recognizing voice amplitude [9]. Male and female voices have different pitch properties [10]. These issues can lead to incorrect speech segmentation. The issues concerning the physical limitations and the privacy of patients probe another challenge in the data acquisition procedures [11].

Effect of speech rate on overall intelligibility in PD is a matter of debate. A comparison of results of previous studies on speech rate in PD is hampered by methodological differences [12].

2. AIM OF THE STUDY

The goal of the review is to investigate inventions in the area of speech disorder assessment applicable to modern mobile devices (MMD). Possible solutions to the problems in real time voice processing through distributed units capable of separating pathological and normal speech signals are targeted. Acoustics features robust to medication changes in Levodopa-responsive patients are identified.

3. SYNTHESIS OF PREVIOUS WORK

In this section, the literature on acoustic analysis of HKD in PD patient have been reviewed and synthesized. Patents related to speech disorder assessment methods and biofeedback devices have been investigated. Choice of publications for review is based on the validation of methodology, credibility of experimental techniques for real-time voice analysis and portability of algorithm to MMDs.

The synthesis is presented chronologically i.e. starting from the investigation of essential acoustic parameters for HKD recognition to the systems implemented for classification between HKD from the normal voice. The review methodology is depicted in Fig. (**1**). Keywords such as "Parkinson disease", "Hypokinetic Dysarthria", "Speech Recognition in real time" were used in the search engines for literature acquisition. The histogram in Fig. (**2**) displays the search engines used, number of total relevant works found and the filtered publications based on the review methodology. The most useful search hits for the synthesis were found using the IEEE Explorer search engine. Other search engines used were Google Scholar, ELIN, EBRARY, LIBRIS and Google Patents.

3.1. Acoustic Measures for HKD Recognition

The investigation of relations between pathological acoustic parameters and Parkinson's disease has been of primary concern for researchers. Acoustic signatures that were robust to phonetic variation in PD patients speech conversations has been identified by Rosen *et al.* [5]. In this experiment, 20 healthy control (HC) and 20 PD patients had 2 minutes of conversational speech. Group differences were detected in T-test in 8 of the 9 measures. This experiment revealed that most specific (95%) and accurate (95%) differentiators of HKD and HC speech were 'Pause Time' and 'Spectral Range'.

In order to characterize the relationship between vocal tract acoustic output and perceptual impressions of pathological speech, another approach has been followed [13]. In this method, voice of vowels 'i', 'u' and 'a' and fricatives 'f' and 's' were of interest during the speech test. Acoustic measures included articulatory rate, segment durations, vowel formant frequencies (formant frequency is the amplitude peak in the sound frequency spectrum), and first moment coefficients. Results indicated high variations in temporal acoustic measures of formant frequency for the PD group when compared to the HC group. These experiments depicted that HKD can be distinguished from HC voice on the basis of acoustic variables such as amplitude variation, spectral range and formant frequency.

Vowel syllables and co-articulation were further investigated to study speech variability between 9 males with PD and a group of 10 HC males [14]. Ratio of F2 formant onset frequency and F2 target frequency was used to infer

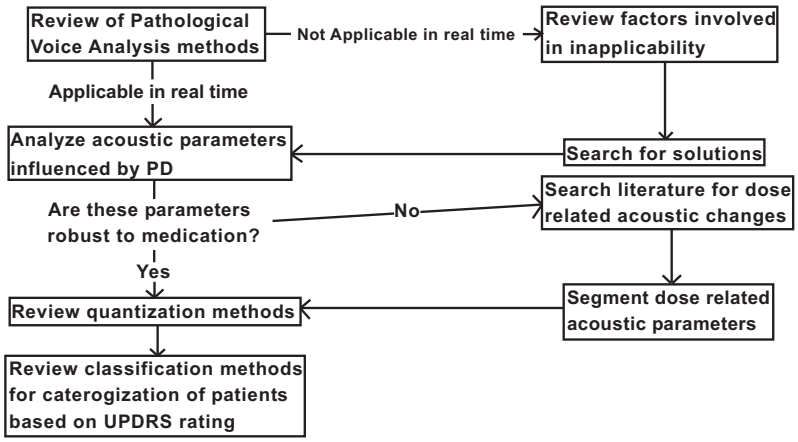


Fig. (1). Review methodology.

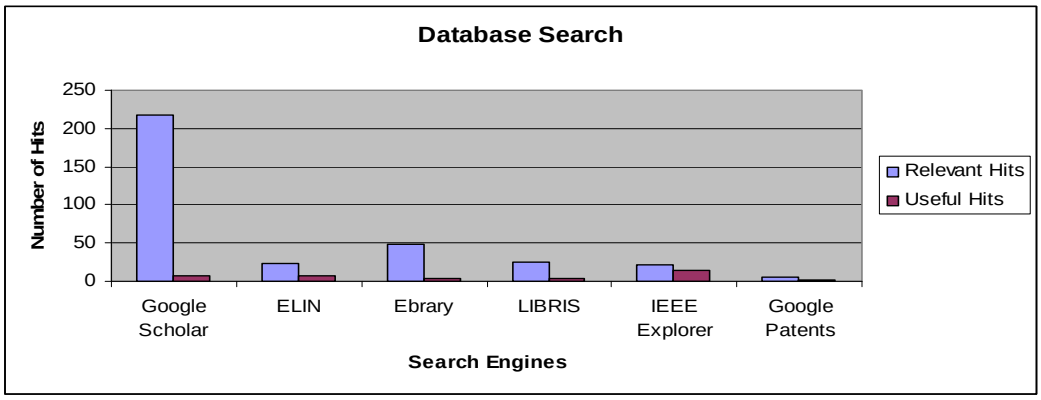


Fig. (2). Database search results.

co-articulation. Fricative F2 was obtained for speech stimuli and compared to F2 onset measures. The result (shown in Fig **3a**, **b**) was formant ratios tended to be smaller for PD speakers than for HC.

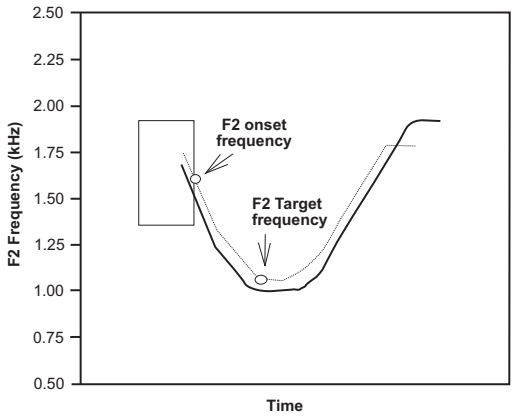
Further investigation of pathological acoustic parameters revealed that for all speaker groups (HC and HKD), STI (Spatiotemporal index, a measure of spatial and temporal variability) values from the loud speech were closest to those from habitual speech [15]. The reason is normalized movement pattern for loud speech resembled that of habitual speech. Further findings [16, 17] depicted that loud speech resulted in intelligibility improvement in HKD acoustics. These studies narrowed down the focus of research to vowel formant time-frequency analysis to classify between pathological and normal voice.

3.2. Vowel Formant Frequency

The most relevant acoustic parameters for production of vowels are the frequencies of first two formants, F1 and F2 [7]. These formant frequencies change as a function of the movements of articulators. In general, the frequency of F2 increases, and that of F1 decreases, as the tongue moves forward (e.g., to form the vowel 'i'), and the frequency of F2 decreases as the tongue moves backward (e.g., to form the vowels 'u' and 'ao'). Also, the frequency of F1 decreases when the tongue is elevated (e.g., to form the vowels 'i' and 'u') and increases when the tongue is lowered, alone or with a downward movement of the jaw (e.g., to form the vowel 'a'). Furthermore, the frequencies of both F1 and F2 decrease when the lips are rounded (e.g., to form the vowel 'u') and increase when the lips are retracted or become unrounded (e.g., to form the vowels 'i' and 'a').

HKD results in vowel formant centralization, that is, formants that have high frequencies tend to have lower frequencies, and formants that have low frequencies tend to have higher frequencies [7]. To improve differentiation of HKD from normal speech, the acoustic metric must be minimally affected by speaker-related variability and maximally affected by the articulatory impairment, as reflected by vowel formant centralization. The metric used was called the formant centralization ratio (FCR). FCR effectively and robustly differentiated the groups with dysarthria from the HC group [7]. Also, the FCR is

(**a**) Two repetitions of word "dice" are shown. Dashed trajectory corresponds to a repetition for PD speaker. Solid trajectory corresponds to repetition for healthy speaker of similar age.



(**b**) Two repetitions of word "dice" illustrate varying co-articulation. Solid repetition was produced at relatively faster speaking rate and has shorter vowel duration; dashed repetition was produced at slower rate and has longer vowel duration.

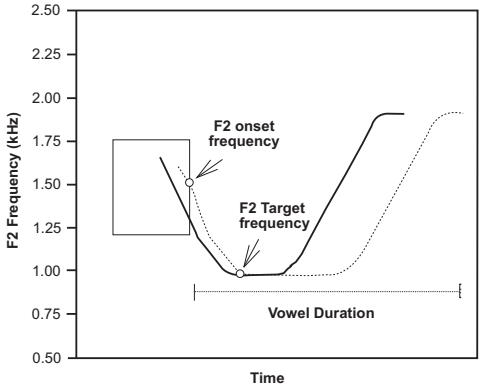


Fig. (3). Formant analysis from Tjaden 2000 [14].

insensitive to gender effects. FCR correlated highly when the correlated variable was the change induced by treatment. These findings suggest that the FCR is a valid and highly sensitive metric of vowel articulation for normal as well as for abnormal speech.

Voice segmentation of speech signals in the case of HKD is difficult because syllable units are spread roughly *via*

intensity changes. For continuous speech segmentation, zero crossing rates (ZCR) provide spectral information at low computational cost. ZCR is the rate when a waveform crosses time axis or changes its algebraic sign. Consonants have high ZCR whereas vowels have low ZCR [18]. Therefore an onset of high ZCR means beginning of consonant and it should be a starting boundary. An offset of high ZCR means an endpoint of consonant and it should be an ending boundary. Segmentation accuracy was drastically improved when ZCR rules were applied to the voice segmentation algorithm [18].

3.3. Automatic HKD Recognition

These studies [13-18] narrowed down the research to formant frequency of consonants and vowels as it contains most relevant acoustics for HKD analysis. The fundamental frequency of voice onset time (VOT) has been used [8] to calculate PD severity ratings (VOT is the interval between the initial articulatory release of a stop consonant and onset of voicing for the subsequent vowel). In this study, the speech datasets were rated from mild to severe. Results indicated that all participants exhibited HKD in OFF medication state. From this experiment, it was observed that Levodopa appeared to have greater effect on VOT within PD group and VOT is a measure of medication related rate change.

Continuous sound power analysis of PD patients for automatic HKD recognition was investigated by Izworski *et al.* [10]. Sound emitted for a long period of time allowed for sound power analysis. Power value P (t) of each frame was calculated by summing up the values of the signal energy within the respective x (a) and x (a+m) limits of the t frame given in equation 1.1 where m is the frame length.

$$P(t) = \sum_{k=a}^{a+m} |x(k)|^2 \tag{1.1}$$

A polynomial p(x) of degree 4 that fits the data, p(x (t)) to P (t) in least squares sense is calculated to represent average values of P (t) is given in equation 1.2.

$$p(x) = p_1 x^n + p_2 x^{n-1} + ... + p_n x + p_{n+1} \tag{1.2}$$

The sum of differences between vector P (t) and p(x) can be used as a voice stability parameter which is given as.

$$Stab = \sum_{k=1}^{a+m} |P(k) - p(k)| \tag{1.3}$$

A comparison between the values obtained from the control group and the patients group is shown in Fig. (**4a**, **b**) respectively. Values of the Stab parameter were much greater for patients than for persons from the control group. With many patients, distinct and varying breaks in phonation amplitude (represented with E (db)) were observed as shown in Fig. (**4b**). With healthy persons, gradual quietening occurred, whereas PD patients ended the emission abruptly. This audio signal was transformed into Fourier series. It was found out that changes were observable especially in the vowels articulation. Results constituted the beginning of tests concentrated on automatic voice classification [10].

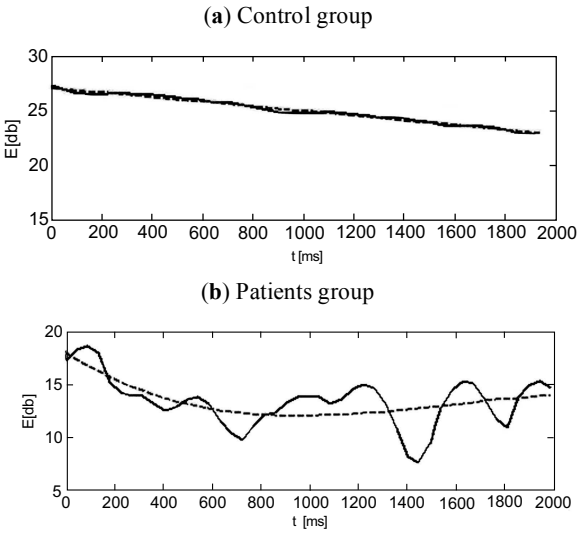


Fig. (4). Sound power analysis from Izworski *et al*. 2006 [10].

In one study of automatic voice classification, artificial neural network (ANN) was used for feature selection and classification of PD voice datasets [19]. Datasets in this study consisted of 195 sustained vowel phonation from 31 people, of which 23 were diagnosed with PD. Adaptive neuro-fuzzy classifier (ANFC) with linguistics hedges gave best recognition results. Pathological voice features such as fundamental frequency of voice, shimmer and jitter were used for classification. ANFC classifier produced 94% accuracy for classification of speech dataset. These results suggest that ANN can be used as a possible tool for automatic HKD classification based on a clinical rating scale as it provides a measure of certainty through supervised learning.

3.4. Automatic HKD Recognition Using Hybrid Approach: Wavelet Analysis and ANN

The voice signal and the related medication effects are functions of time. Therefore it is preferable to quantize the speech datasets in time and frequency scales before the speech data is classified. Linear predictive coding (LPC) has been used [20] to quantize PD patients' voice based on the formants F1, F2 and F3. Comparing with the normal voice, pathological voice showed high variations. Although LPC distinguished HKD from normal voice, it does not give adequate quantification values to make the decision. A hybrid approach using wavelet analysis and ANN was introduced [20]. Normalized energies and entropies of wavelet coefficients were used to formulate feature vector of speech sample. This feature vector was used to identify HKD. A 3-layered feed forward ANN with back propagation (BP) algorithm was used for classification. The flowchart of this algorithm is shown in Fig. (**5**).

A total of 100 pronounced words were used for classification out of which 50 words were recorded from the normal speakers and the other 50 were recorded from the patients with HKD. A set of 80 words (40 normal and 40 pathological) was used for training ANN, whereas a set of 20 words (10 normal and 10 pathological) was used for testing ANN. Using Discrete Wavelets Transforms (DWT) a clear difference was noticed between wavelet evolutions of HKD

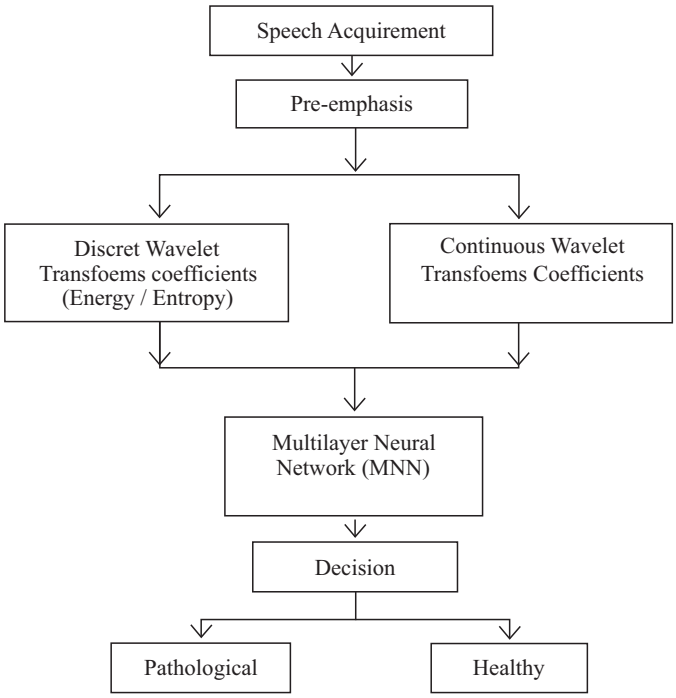


Fig. (5). A hybrid approach using wavelet transform and ANN for HKD recognition from Salhi *et al.* 2010 [20].

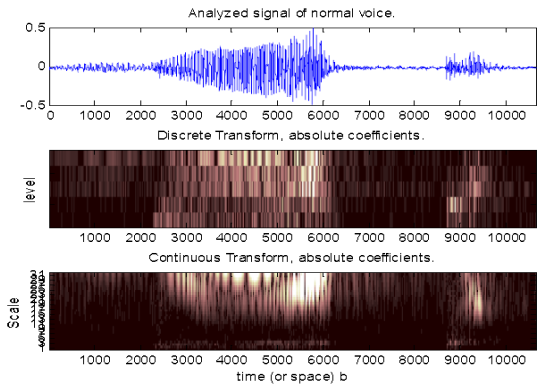
and normal voice. A visual pattern is shown in Fig. (**6a**, **b**). ANN classification with five DWT coefficients produced 90% classification rate for pathological voice and 100% classification rate for normal voice [11].

An analysis showed that the formant analysis methods using Fourier Transform is not a suitable tool for speech analysis [11]. Since speech is a highly non-stationary signal, wavelet transform proved to be a better tool for analysis of non stationary signals [20] as it is useful in localizing a symptom both in time and frequency scales.

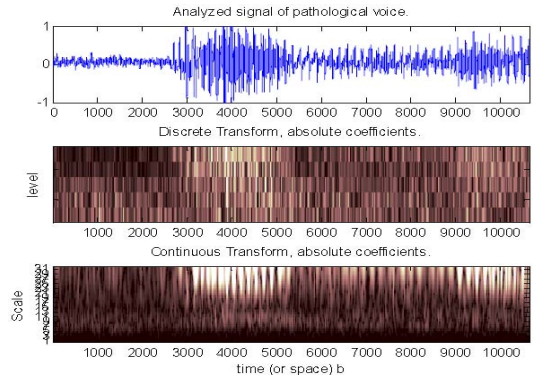
3.5. Review of Patents

The following review of patents focuses on real-time speech assessment methods in biofeedback devices and distributed speech processing mobile units. Speech analysis methods used in speech therapy techniques has also been evaluated for their possible use in automatic HKD recognition.

A biofeedback system for speech disorders has been reported by Kehoe and Sereno [21]. This system could detect disfluent speech and provide auditory feedback to the speaker, enabling him to produce immediate fluent speech. This system could shift pitch of the speaker's voice according to speaker's disfluent speech using an electromyography (EMG). The biofeedback system also supported user's control of speaking rate using a Masking Auditory Feedback (MAF). The MAF improved speaker's awareness of physical sensation of speech. This invention could only distinguish between healthy and pathological speech, but it probed the research in biofeedback systems for speech disorder analysis.



a. Normal Voice



b. Pathological Voice

Fig. (6). Wavelet analysis from Salhi *et al.* 2010 [11].

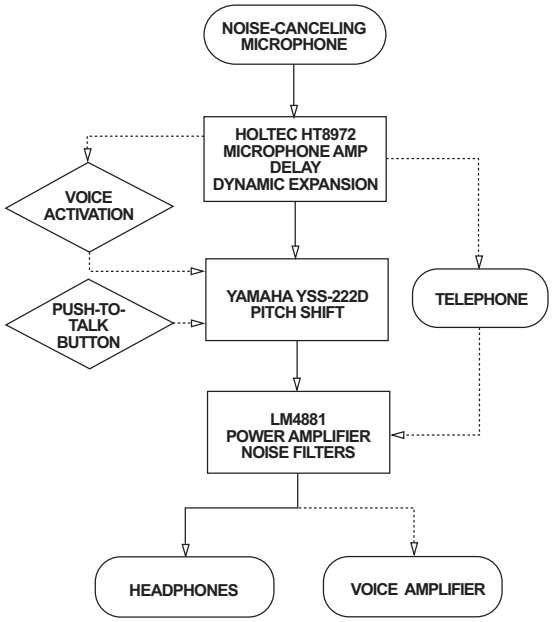


Fig. (7). A Speech aid biofeedback system for HKD patients [22].

Another biofeedback system to aid persons with HKD was reported by Kehoe [22]. This system could produce multi talker babble noise to the speaker's ear to prompt him to speak louder. The invention used combination of Delayed Auditory Feedback (DAF) and Altered Auditory Feedback (FAF) for pitch and frequency shifting to induce PD patients to speak clearly. In this way, intelligibility in the user's speech could be improved. This system may be used with a microphone or headphones with a voice amplifier or it may be connected to a telephone as shown in Fig. (**7**). The benefit of using this system was that no mental stress or training was needed which was beneficial for patients with cognitive disorders.

A speech-based biofeedback system was reported by Ron [23]. This system could monitor subjects' emotional state using an onboard digital signal processor to analyze speech parameters associated with emotions. Information about emotional state was carried out at the remote end vocally using a telephone line or textually through the internet for further processing. This invention could particularly be useful for mood variability analysis in PD patients [24].

Speech disorder assessment may be done by performing a comparative analysis between the user pronunciation and the actual pronunciation recorded in a database [25-28]. In these systems, users were asked to pronounce a given word through a microphone attached with a processor. The user speech pronunciation patterns were compared with the ideal speech pronunciation patterns stored in the database to determine if there were any errors in the pronunciations. These errors were categorized depending upon therapists' scoring of the user's pronunciations. Based on these scores speech disorder severities could be determined. A similar system based on comparative analysis for speech assessment of aphasia patients has been presented [29], which may be distributed across multiple devices or computer platforms (Fig. **8**).

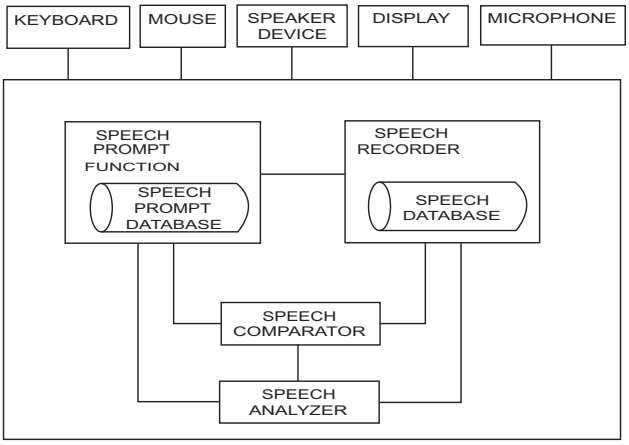


Fig. (8). A system for the speech assessment of aphasia patients [29].

A distributed speech recognition model is reported by Reding *et al.* [30]. In this method, speech was recorded using a microphone and transmitted to speech processing facility over the internet. The facility provided speech recognition services on the digitized speech. In the case of real time speech recognition for voice dialing, the size of vocabulary is limited due to the limited processing power of portable devices. Using this method, the speech processing facility may augment a device incapable of speech processing with voice recognition capability. Another invention on remote assessment of speech is reported by Moran *et al.* [31]. In this invention the application software was resided in a server. This software comprised of a feature extraction engine (FE) and a comparator (CMP) as shown in Fig. (**9**). The server could interact across the network with a client device to obtain speech datasets from users. These speech datasets were stored in a database. The feature extraction engine was used to extract one or more acoustic features from the speech signals stored in the database. These features were compared with the features of referenced speech samples to assess the speech disorder. Notice that speech disorder assessment algorithms are computationally expensive and can be very time consuming with limited processing units. These inventions direct a possibility of HKD assessment in PD over the internet using a mobile device connected to a centralized database server.

The so-called Lee Silverman voice treatment (LSVT) therapy system was introduced for speech and movement disorders in a patent by Ramig *et al.* [32]. The LSVT consisted of a variety of voice exercises including sustained vowel phonation, pitch exercises, reading and conversational activities. This speech therapy was used to improve speech impairment in PD patients as their speech deteriorates with the disease progression. Therapy can be provided alone or in conjunction with medication like Levodopa. An extension of this work was made by embedding LSVT therapy system in a mobile device known as LSVT Companion (LSVTC) [33]. LSVTC was programmed to collect data on sound pressure level (SPL), fundamental frequency (F0) and duration of phonation. It was used to provide feedback to individuals on their performance during LSVT therapy. LSVTC was

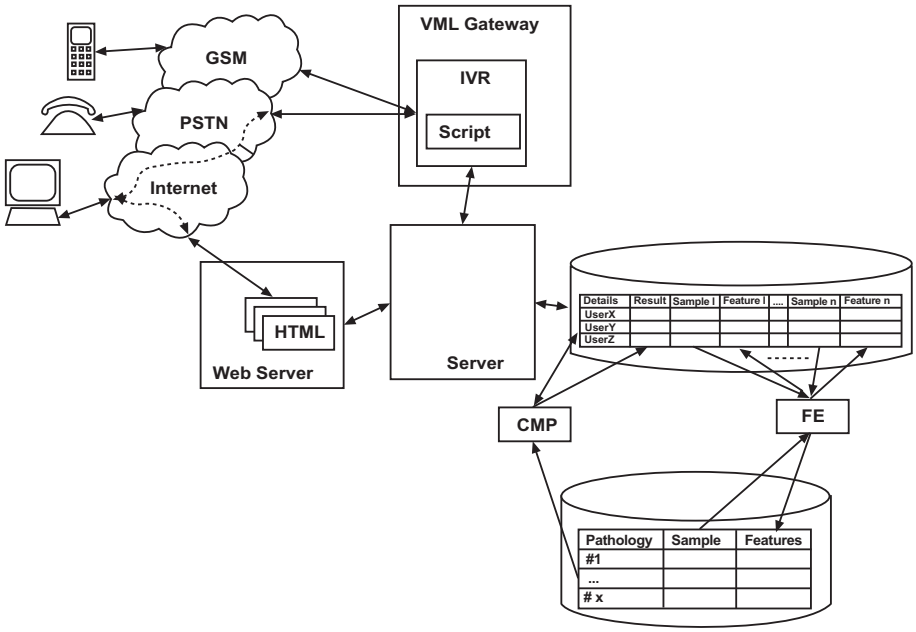


Fig. (9). Telephonic pathology assessment by a comparative analysis between pathological and normal speech patterns [31].

employed with simple bar graphs to indicate SPL, pitch, and time as shown in Fig (**10a**). Fig. (**10b**) shows this graph when applied to conversation exercise. Using colored boundaries, patients could maintain the sound pressure level during their voice therapy. A user interface was developed in LSVTC to receive feedback for sustained vowel duration exercise is shown in Fig. (**10c**). In this interface SPL was expressed by a thermometer, where the blue "mercury column" was used to raise according to the instant loudness. The colored bands on the back plate indicated the targets. Using this information, the patients could accommodate the duration of speech and sound pressure level during the therapy which could improve the respiratory and articulatory system of the patient.

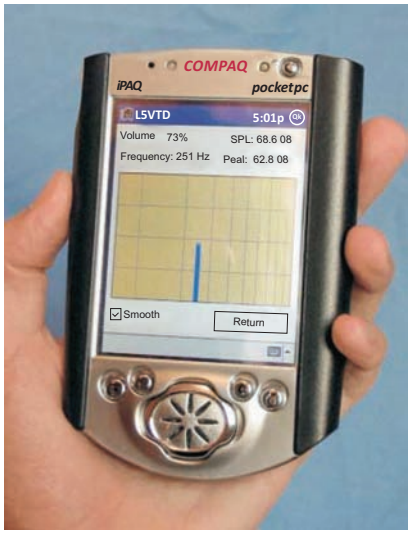


Fig. (10a). LSVTC with simple bar to indicate SPL and pitch from Matos *et al*. [33].

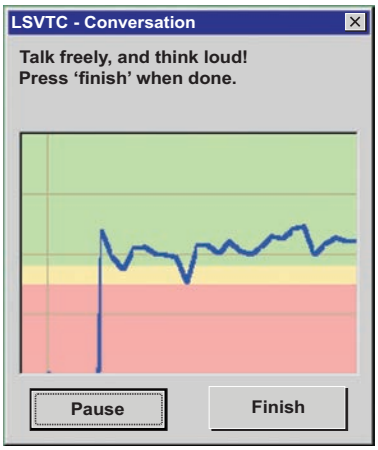


Fig. (10b). Screenshot of a graph applied to the conversation exercise [33].

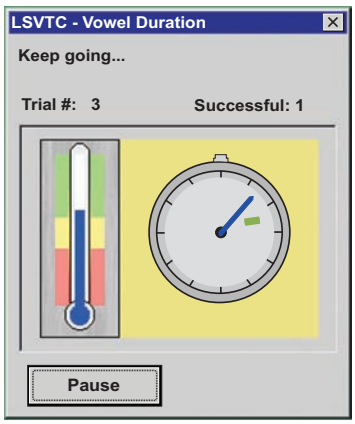


Fig. (10c). Way to receive feedback for sustained vowel duration exercise [33].

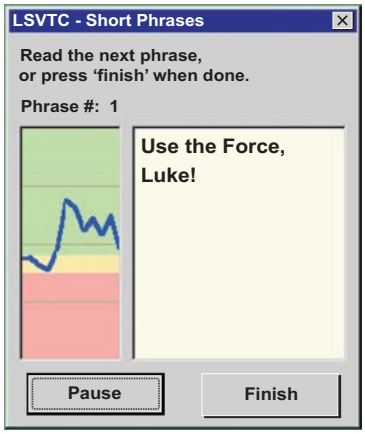


Fig. (10d). Screenshot of a graph applied to functional phrases exercise [33].

In this review of patents, some inventions of biofeedback systems for speech assessment and distributed sound processing units are described. In the reported biofeedback systems, the speech aid is provided using frequency and pitch shifting. In other patents, the speech assessment is reported to have been done by a comparative analysis between user speech patterns and referenced speech patterns stored in a database. None of the patents were reported to have used voice quantization methods (e.g. wavelet transform described in section 3.4) for pathological speech assessment. No patents were further incorporating a method for pathological voice quantification based on a clinical rating scale. The review of patents on distributed speech processing units though directs a possibility of using a mobile device to assess pathological speech.

4. CURRENT AND FUTURE DEVELOPMENTS

A major challenge in HKD assessment through mobile devices is the voice segmentation and noise removal in speech signals acquired in the noisy environment. An automatic processor should succeed at identifying silence, vocalic, fricative and nasal sounds and stop gaps to provide gains in speech intelligibility. Especially in case of HKD, voice segmentation is difficult because syllable units are spread roughly *via* intensity changes but exact boundary positions are elusive in successive vowels.

Currently researchers are focused at achieving higher reliability in speech assessment systems because voice recording and assessment may differ in uncontrolled acoustic environments. Also acoustic features like sound-pressure level and formant analysis are gender dependent. To deal with uncontrollable noise effects, a measure of dysphonia called pitch period entropy (PPE) is introduced for HKD analysis in a recent work [34]. PPE along with other acoustic features (jitter, shimmer, and Noise-to-Harmonic ratio) was used to train a SVM (support vector machine). Results showed that PPE is robust to perplexed acoustic changes and gender voice.

In future, research and development targeted to produce an automatic HKD assessment algorithm that is embedded in a mobile device is advocated. The advanced sound processing capabilities in the mobile companion allows the feasibility of auditory and visualized feedback to the patients (an example is the LSVT companion). Medicine adjustments could possibly be made based on the speech symptom severity ratings thus supporting timely treatment of PD patients in the home environment.

In conclusion, this review evaluated HKD recognition methods for PD speech symptom severity assessment using datasets from mobile devices. Investigation of patents filed for speech disorder assessment revealed that a comparative analysis of abnormal and normal speech can be made efficiently in distributed speech processing units. Vowel and consonant formants are the most relevant acoustic parameters to reflect PD medication changes. Since voice signals are highly non-stationary, focus was laid on quantization of speech signals. It was found that time-frequency methods fail to quantify voice signals over time-series based on voice frequency. Wavelet transform has the potential to quantize the non-stationary voice signals over the time-series using scale and translation parameters. In this way voice intelligibility in the waveforms can be analyzed in each time frame. Artificial neural network and support vector machines are possibly useful techniques for voice characterization. Patients' speech can be categorized according to clinical speech ratings. It is concluded that modern mobile devices could prove to be efficient tools to monitor patient's speech symptoms on a daily basis.

ACKNOWLEGDEMENT

This research is a module of the project "E-Motions". This project has been running in Dalarna University in collaboration with Abbott Laboratories and Nordforce Technology and is funded by a grant from Swedish Knowledge Foundation.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

REFERENCES

[1] S.A. Factor, W.J. Weiner, "Parkinson's disease: Diagnosis and clinical management", New York: Demos Medical Publishing 2002.

[2] L. Hartelius, P. Svensson, "Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: A survey", *Folia Phoniatr et Logop*., vol. 46: pp. 9-17, 1996.

[3] J.E. Zicker, W.J. Tompkins, R.T. Rubow, H.A. James, A portable micro- processor-based biofeedback training device", *IEEE Trans Biomed Eng*., vol. 27(9): pp. 509-15, 1980.

[4] K.Y. Joan, T.L. Whitehill, S.Y.S. Susanne, "Intonation contrast in cantonese speakers with hypokinetic dysarthria associated with Parkinson's disease", *J Speech Lang Hear Res*., vol. 53: pp. 836-49, 2010.

[5] K.M. Rosen, R.D. Kent, A.L. Delaney, "Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers", *J Speech Lang Hear Res*., vol. 49: pp. 395-411, 2006.

[6] J. Camburn, S. Countryman, J. Schwantz, "*Parkinson's disease: Speaking out*", Denver, CO: The National Parkinson Foundation 1998.

[7] S. Sapir, L.O. Ramig, J.L. Spielman, C. Fox, "Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech", *J. Speech Lang Hear Res*., vol. 53: pp. 114-25, 2010.

[8] E. Budkowski, "Voice onset time in Parkinson disease", MS thesis, Bowling Green State University, Ohio, USA, 2007.

[9] A.C. Dominguez, "Pre-processing of speech signals for noisy and band-limited channels", MS thesis. KTH, Stockholm, Sweden 2009.

[10] A. Izworski, P. Augustyniak, T. Orzechowski, "Processing and analysis of voice anomalies in course of Parkinson's diseases", *Proceedings of Eighth IASTED International Conference.* Honolulu, Hawaii, USA, 2006: pp. 354-57.

[11] L. Salhi, T. Mourad, A. Cherif, "Voice disorders identification using multilayer neural network", *IAJIT*, vol. 7: pp. 177-85, 2010.

[12] D.M. Letter, P. Santens, D.M. Bodt, B.J. Borsel, "Levodopa induced alterations in speech rate in advanced Parkinson's disease", *Acta neurol. Belg, vol.* 106: pp. 19-22, 2006.

[13] P. McRae, K. Tjaden, "Acoustic and perceptual consequences of articulatory rate change in Parkinson disease", *J. Speech Lang Hear Res*., vol. 45: pp. 35-50, 2002.

[14] K. Tjaden, "An acoustic study of Co articulation in dysarthric speakers with Parkinson disease", *J Speech Lang Hear Res*., vol. 43: pp. 1466-80, 2000.

[15] J. Kleinow, A. Smith, L.O. Ramig, "Speech motor stability in IPD: Effects of rate and loudness manipulations", *J Speech* Lang Hear Res., 2001, vol. 44: pp. 1041-51, 2001.

[16] A.T. Neel, "Effects of loud and amplified speech on sentence and word intelligibility in Parkinson disease", *J. Speech Lang Hear Res.*, vol. 52: pp. 1021-33, 2009.

[17] K. Tjaden, G.E. Wilding, "Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings", *J. Speech Lang Hear Res*., vol. 47: pp. 766-83, 2004.

[18] X. Zhao, D. O'Shaughnessy, "A new hybrid approach for automatic speech signal segmentation using silence signal detection, energy convex hull and spectral variation", *In Proceedings of Canadian Conference on Electrical and Computer Engineering*, Ottawa, USA, 2008: 145-8.

[19] M.F. Caglar, B. Cetisli, I.B. Toprak, "Automatic recognition of Parkinson's disease from sustained phonation tests using ANN and adaptive neuro-fuzzy classifier", *J Eng Sci Design*, vol. 1: pp. 59-64, 2010.

[20] L. Salhi, M. Talbi, A. Cherif, "Voice disorders identification using hybrid approach: Wavelet analysis and multilayer neural networks", WASET, vol. 45: pp. 330-9, 2008.

[21] T.D. Kehoe, M. Sereno, *"Biofeedback system for speech disorders"*, U.S. Patent 5794203, 1998.

[22] T.D. Kehoe, *"Electronic speech aid and method for use thereof to treat hypokinetic dysarthria"*, U.S. Patent 20100100388, 2010.

[23] S. Ron, *"Speech-based biofeedback method and system"*, U.S. Patent 5647834, 1997.

[24] K. Shifren, K. Hooker, P. Wood, J.R. Nesselroade, "Structure and variation of mood in individuals with parkinson's disease: A dynamic factor analysis", *Psych Aging*, vol. 12: pp. 328-39, 1997.

[25] C. Waryas, J.H. Segapeli, A.V. Spiser, "*Speech analysis system and method*", U.S. Patent 6725198, 2004.

[26] R. Bogdashevsky, V. Alexeev, V. Yarigin, G. Baker, H. Stanton, *"Speech signal processing for determining psychological or physiological characterstics using a knowledge base"*, U.S. Patent 6006188, 1999.

[27] S.K. Gupta, P. Raghavan, C. Vincchi, "*Automatic assessment of phonological processes*", U.S. Patent 7302389, 2007.

[28] S.G. Fletcher, B. Faber, "*Real time voice analysis and method for providing speech therapy*", U.S. Patent 20070168187, 2007.

[29] K.L. Haley, "*Methods, systems and computer program products for speech assessment"*, U.S. Patent 20090275005, 2009.

[30] C. Reding, S. Levas, "*Methods and apparatus for performing speech recognition and using speech recognition results*", U.S. Patent 6915262, 2005.

[31] R. Moran, R. Reilly, D. Chazal, O. Mullane, P. Lacy, "*Telephone pathology assessment*", U.S. Patent 7457753, 2008.

[32] L.O. Ramig, C.M. Fox, D. McFarland, B.G. Farley, "*Total communications and body therapy*", U.S. Patent 7762264, 2010.

[33] C. Matos, A. Halpern, L. Ramig, J. Spielman, J. Bennet, "*Updates to PDA-enhanced speech treatment for Parkinson's disease"*, Available at: https://www.cu.edu/ColemanInstitute/Boulder-material/posters/Matos.pdf (Accessed on: March 18, 2011).

[34] M.A. Little, E.P. McSharry, J. Spielman, L.O. Ramig, "Suitability of dysphonia measurements for tele-monitoring of Parkinson's disease", *IEEE Trans Biomed Eng*., vol. 56: pp. 1015-22, 2009.