Audio Engineering Society

Convention Paper

Presented at the 123rd Convention
2007 October 5–8 New York, NY, USA

Special Hearing Aid for Stuttering People

Piotr Odya[1], Andrzej Czyżewski[1,2]

[1] Gdansk University of Technology, Multimedia Systems Department, 80-952 Gdansk, Poland
[2] Excellence Center PROKSIM (International Center of Hearing and Speech of the Institute of Physiology and Pathology of Hearing), 01-943 Warsaw, Poland
piotrod@sound.eti.pg.gda.pl
ac@pg.gda.pl

ABSTRACT

Owing to recent progress in digital signal processors developments it has been possible to build a subminiature device combining speech and hearing aid. Furthermore, despite its small dimensions, the device can execute quite complex algorithms and can be easily reprogrammed. The paper puts an emphasis on issues related to the design and implementation of algorithms applicable to both speech and hearing aids. Frequency shifting or delaying the audio signal are often used for speech fluency improvement. The basic frequency altering algorithm (FAF) is similar to the sound compression algorithm used in some special hearing aid as above. Therefore, the experimental device presented in the paper provides a universal hearing & speech aid which may be used by hearing or by speech impaired persons or by persons suffering from both problems, simultaneously.

1. INTRODUCTION

Modern Digital Signal Processors (DSP) may have small dimensions and very low current consumption, but they are able to execute complex algorithms. In addition, they can be easily reprogrammed using a standard PC computer. Taking advantage of these processors, it was possible to build a device, which can be used either as a speech aid or a hearing aid, or both.

In various countries the number of stuttering persons ranges from approximately 0.5% to 1% of the population. Considering the fact that typically every fourth case of stuttering is a medium-advanced and requires therapy, one can estimate that in average sized countries the demand for the device might reach several dozen thousand persons. In recent years, in the Multimedia Systems Department research studies have been conducted to develop an electronic device for stuttering people. The first version has dimensions of a cigarette package and required the use of external microphone and headphones [4] [10] [21] [23]. The new version of the device has size of a regular ear-canal hearing aid. Consequently, the patients can easily use it in everyday situations. Interlocutors might not even realize their conversation partner uses a speech corrector. The device is called the Subminiature Digital Speech Aid (SDSA).

On the basis of the SDSA hardware, it was possible to build also other kind of devices, e.g. special hearing aid. Recent screening hearing tests, which have been carried out in Poland, showed that many people suffer from hearing loss [5] [6]. Worse still, typical hearing aids are not able to help some particular groups of patients, e.g. newborn infants, people working in a noisy environment, aircraft pilots or patients with cochlear implants. Owing to hybrid signal processors, it was possible to implement algorithms of spectral transposition. The aim was to amplify essential components of speech signal and shifting them in a low frequency band, and at the same time filtering out the background noise.

2. SUBMINIATURE DIGITAL SPEECH AID

There are many theories concerning stuttering causes [1] [14] [22] [24] [25]. Following the successful introduction of the first compact digital anti-stuttering device [4] [23] long-term research studies were carried out in the Multimedia Systems Department and in the International Center of Hearing and Speech of the Institute of Physiology and Pathology of Hearing. The results of this study proves that the main reason of stuttering should be sought within the auditory feedback loop disorders. Moreover, it was confirmed empirically that by introducing some changes into the auditory feedback loop a serious reduction of stuttering may be achieved. Even though the background of this phenomenon was not explained yet, alterations to the auditory feedback loop were being introduced for many years. There are various methods existing of changing speech in the auditory feedback loop. All this methods change the way a patient perceives her or his own speech. The most popular methods are DAF (Delayed Auditory feedback) and FAF (Frequency Altered Feedback) [10] [13] [15] [16] [17] [19] [20] [24].

2.1. Tests with the DSA

The first stage of this study was to assess various speech correction algorithms and their efficiency. For this purpose the Digital Speech Aid (DSA) has been employed. As mentioned before, the DSA was designed in early 90s at the Gdansk University of Technology in co-operation with the Canadian Dalhousie University. The DSA allowed using the DAF and FAF methods simultaneously or separately. The satisfactory results in clinical trials the Multimedia Systems Department of the Gdansk University of Technology convinced the Foundation for Polish Science to subsidize testing of 100 units of the DSA within the program of supporting pre-commercial works on new technologies, products and services. In this way, 100 devices were under examination in 100 selected psychological-pedagogical centers in various locations in Poland (see black dots in Figure 1). Comprehensive assessments of the efficacy of the therapy using DSA have been carried out using questionnaires and a so- called "syllable test" [21].

Figure 1: Map of Poland with indicated places where tests have been carried out.

128 patients took part in the experiments, most of them were school age children. 75% of patients were boys – this confirms numerous observations that males more often stutter than females. Merely 30% of patients had a stuttering person in a family, this indicates that there is no direct relationship between stuttering and familial inheritance patterns.

In-depth analysis of results obtained from the Centers, performed in the Statistica software, showed that improvement of total fluency was observed in above 75% of stutterers. Furthermore, in more than 70% cases speech remained fluent even after switching off the DSA. FAF and DAF+FAF methods have been assessed as more efficient than while using the DAF method only (Figure 2). Also the syllable tests proved the high efficiency of spectral alteration methods. In a case of the FAF method, patients most often have chosen a spectral transposition down the frequency scale which amounted to 6% of the octave.

Based on the recorded patients' utterances it was possible to perform a thorough speech analysis. Emphasis was placed on observation of a vocal tone frequency, and formant frequencies and their amplitudes. Modified cepstral analysis was performed, employing methods developed in the Multimedia Systems Department [3] [10] [18].

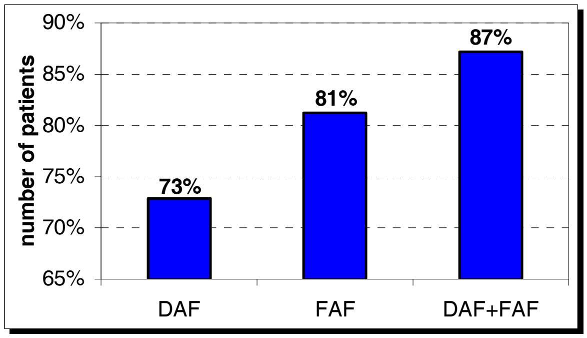

Figure 2: Number of patients who have noticed speech fluency improvement.

The vocal tone frequency was evaluated according to the block diagram presented in Figure 3. The aim of the spectrum normalization block is to enhance components responsible for the vocal tone, especially in noise presence. Cepstral Coefficients were computed using the following Equation:

$$C_r = \sum_{i=1}^{m} N_i \cos\left(r\frac{i\pi}{m}\right) \qquad m = \frac{l_{pr}}{f_p} f_c \tag{1}$$

where: r – order of cepstral coefficient; i – number of consecutive spectral component; N_i – normalized value of the spectrum logarithm; l_{pr} – number of sample in a segment; f_p – sampling frequency; f_c – max. frequency of cepstrum analysis.

Cepstrum smoothing was performed according to the expression:

$$W_n = \sum_{r=1}^{r_{\max}} C_r \cos\left(r\frac{n\pi}{m}\right) \tag{2}$$

where n is smoothing order.

In the next step, the vocal tone frequency was estimated:

$$\hat{f} = \frac{1}{r_c} \cdot \frac{\sum_{i=m}^{n} W_i}{\sum_{i=m}^{n} i \cdot W_i} \tag{3}$$

where: r_c cepstral analysis resolution; W_i – coefficient number i; $m=k-1$, $n=k+1$ – where k is the number of the maximal cepstral coefficient (related to the vocal tone).

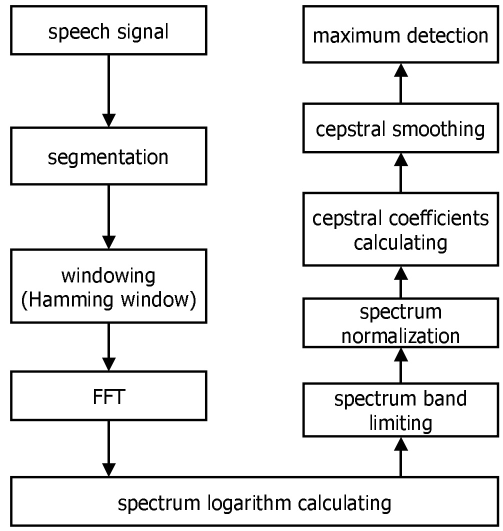

Figure 3: Block diagram of the vocal tone frequency estimation algorithm.

The algorithm for detecting formants frequency and amplitude is similar to the described above. The main differences consist in absence of the spectrum normalization block. Additionally, the preemphasis (6dB/oct.) is employed before the segmentation. In this way, formants detection is easier.

The algorithms were implemented in the Matlab environment. Analysis parameters were as follows:

- sampling frequency – 22050Hz;
- frame length – 1024 samples
- overlap length – 583 samples;
- the band width – 2756.25Hz (vocal tone estimation) and 5512.5Hz (formants detection);
- smoothing order – 66 (vocal tone estimation) and 12 (formants detection).

The analysis has been performed on recordings sent from the psychological-pedagogical centers. Examples of results are presented in the Tables 1-3.

patient	vowel	F_0 [Hz]	
		without DSA	DAF
R23	a	214.42	210.63
R23	a	305.36	226.77
R23	e	256.00	196.17
R23	o	210.79	241.83
R23	o	236.33	209.37
R23	u	293.95	300.72
R34	a	271.40	245.06
R34	o	215.28	240.65

Table 1 Examples of vocal tone frequencies – DAF method

patient	vowel	F_0 [Hz]	
		without DSA	FAF
R19	a	250.80	270.10
R19	o	240.37	256.20
R80	a	232.74	278.20
R80	o	272.93	259.96
R80	u	239.23	281.81
R80	y	139.12	229.48

Table 2 Examples of vocal tone frequencies – FAF method

patient	vowel	F_0 [Hz]	
		without DSA	DAF+FAF
R51	e	264.82	268.13
R51	i	270.95	229.37
R51	o	260.04	261.55
R65	a	241.53	236.50
R65	e	221.18	234.53
R65	o	257.80	254.25
R65	u	296.02	279.45

Table 3 Examples of vocal tone frequencies – DAF+FAF method

The analysis showed that the FAF method causes significant ($p<0.07$, independent t-test) rising in the vocal tone frequency. This phenomenon is caused by attempts at compensating for the perceived pitch-shift [11] [12]. The DAF methods caused lowering the vocal tone frequency as a result of muscle tension reducing while talking with the DSA. In the case of two methods employed (DAF+FAF), significant changes were not observed. Probably, adding some delay resulted in reducing correlation between articulated and perceived speech and the compensation did not occur. Changes in the formants frequency and amplitude were not significant and they could not be generalized.

The tests results indicated also that some improvements in the DSA construction should be introduced. For example it occurs that in most cases there exists a dominant ear in stuttering person, to which the altered sound could be provided. In this case it is sufficient for this patient to modify the feedback loop monoaurally to reduce stuttering. Also, many patients were embarrassed to use the DSA because of its visible casing and earphones. That is why it was decided to develop a new version, the Subminiature Digital Speech Aid (SDSA).

2.2. SDSA Prototype

A prototype of the new version of the device has been built (SDSA) that should provide a comfort to stutters and - at the same time – it is easy to operate. While using the SDSA only one headphone is required. It means that the altered speech is provided only to one ear. The second ear can receive non-altered sounds, which helps to perceive the interlocutor's speech.

The Toccata Plus digital signal processor was chosen as a core part of the device engineered. The main advantages of the Toccata processor are: low power consumption (typically about 400μA at 1.28MHz clock), small size (5.97x3.48x1.52 mm with built-in EEPROM) and programming flexibility (it can be easily reprogrammed using a PC computer). The processor has two subsystems that operate concurrently: the RCore, a fully programmable DSP core, and the Weighted Overlap-Add (WOLA) Filterbank coprocessor, a dedicated, configurable processor that transforms signals to the time-frequency domain [26]. Typical applications of the Toccata processor are hearing aids, but since it is programmable, it can be used in various applications [2] [8].

Parameters of the Toccata processor are set as follows: system clock frequency: 1.92MHz, sampling frequency: 16kHz, subbands number: 16 (32 points FFT). Measured current consumption was typically equal to about 300μA, maximum – 800μA.

The block diagram of the SDSA is presented in Figure 4, and the prototype in Figure 5.

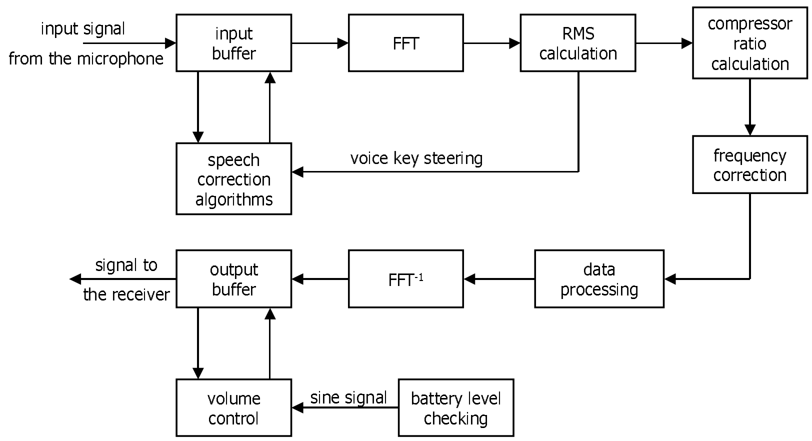

Figure 4: Block diagram of the SDSA.

Figure 5: Prototype of the SDSA.

The SDSA implements some well-known auditory feedback loop signal processing algorithms, such as delaying (DAF) or frequency altering (FAF) speech signal, but also enables to use combination of these methods (DAF and FAF, FAF-to-DAF morphing). Moreover, new algorithms such as: reverb or delay modulation, not used before in the anti-stuttering devices, have been implemented on this processor. Some additional algorithms for sound enhancement are also utilized, e.g. dynamic processor, equalizer (16 bands), battery monitor and voice key. The aforementioned algorithm was well tolerated by the stuttering person, especially that it causes the SDSA to operate only when the person is speaking, which is very desirable in every day usage.

Algorithms and their parameters can be changed according to patients needs employing special software. The software work under Microsoft Windows based operating systems was developed (along with an appropriate PC interface) within the framework of the project. In Figure 6 the main window of the software is presented.

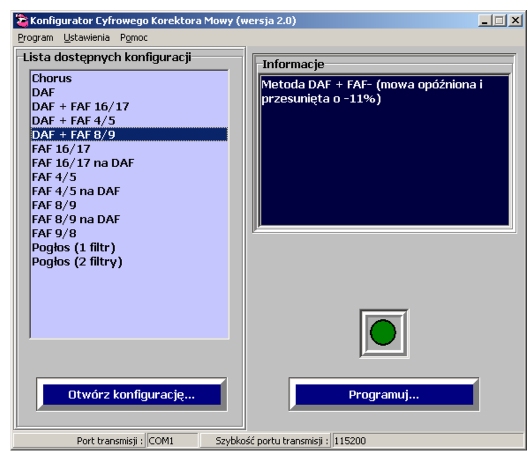

Figure 6: User interface of the developed software.

2.3. Test series

Three male patients took part in the tests. They were asked to read some short texts excerpts (about 200 syllables) with and without the SDSA and speech correction. All speech disfluency occurrences were counted and then the relative error was calculated as the ratio of the number of disfluency occurrences to the number of syllables.

The algorithms parameters were as follows:

- DAF – delay 70ms;
- FAF – 6% octave transposition down the frequency scale;

- DAF+FAF – delay 70ms, 6% octave transposition down the frequency scale;
- reverb algorithm – reflection times: 85 and 103ms;
- delay modulation I – period: 2s, average delay: 64ms, delay range: +/-40ms, modulation function: sinusoidal;
- delay modulation II – period: 3.2s, average delay: 50ms, delay range: +/-30ms, modulation function: sinusoidal and saw wave combination;
- FAF-to-DAF morphing – delay 70ms, 6% octave trans
- position down the frequency scale.

All the methods tested turned out to be efficient and most of them improved speech fluency significantly. The most efficient methods are DAF+FAF, FAF into DAF and delay modulation II (Figure 7).

Similarly to the preliminary tests with the DSA, speech analyses have been carried out. The results indicated that there are significant differences in a vocal tone frequency in comparison to the DSA. First of all the vocal tone frequency changes depended more on a patient than the alteration method used. This is caused by the fact that the second ear receives unprocessed sounds. This means that even though changes to the auditory feedback loop are introduced and as a result speech fluency is observed, but the vocal tone frequency changes are of individual nature and depend on a patient. The results are presented in Table 4. In the table *p* values (for the independent t-test) are given. The background color of the cells of this table indicates a trend of the vocal tone frequency changes. Black background indicates the vocal tone frequency lowering, grey – rising. Detailed results are not presented here since they are too numerous.

3. HEARING AID

Some experiments pointed out that moving the essential components of speech signal into the low frequency band can improve speech intelligibility in some cases of hearing loss. This concerns mostly hearing impaired persons in which a so-called residual hearing is retained. It is caused by the fact that such persons maintain

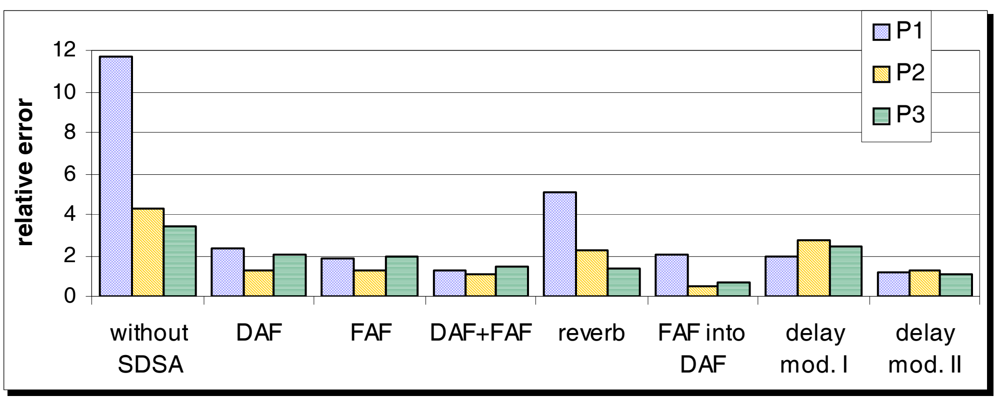

Figure 7: Efficiency of various speech correction methods.

patient	DAF	FAF	DAF+FAF	reverb	FAF into DAF	delay mod. I	delay mod. II
P1	0.027	0.768	0.375	0.350	0.377	0.878	0.966
P2	0.027	0.738	0.137	0.072	0.779	0.022	0.142
P3	0.218	0.864	0.084	0.687	0.753	0.031	0.454
all	0.773	0.677	0.392	0.167	0.571	0.109	0.358

Table 4 Analysis results,
the *p* values for various methods and individual patients are given in cells.

hearing ability in the range of low frequencies (Figure 8) [7] [9]. To help these patients, the audio signal is proportionally transposed down the frequency scale by dividing each frequency component by a factor. This process compresses the speech spectrum in order to introduce as much information as possible into the limited audible frequency range of the hearing impaired listener. Two different algorithms are proposed for such purpose. The first algorithm employs a special resampling technique similarly to the FAF method. It gives better speech quality, but has some limitations (amount of the shifting down the frequency scale cannot exceed about 50%). The second algorithm operates in the frequency domain, its advantage is larger amount of the shifting to obtain. Unfortunately, the speech resulting from this has worse quality and sounds unnaturally. Formant pattern changes are the main reason for such a phenomenon.

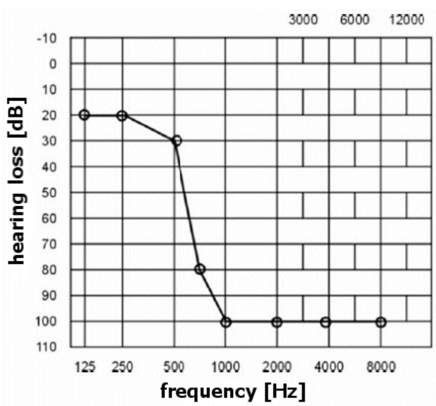

Figure 8: An example of the residual hearing loss (corner audiogram).

Preliminary experiments showed that transposing the speech frequency can help patients with a hearing loss to understand speech, but in most cases habituation is required to understand the transposed speech, especially while using the latter mentioned method. At the moment the device is under a clinical trial at the Institute of Physiology and Pathology of Hearing in Warsaw.

4. CONCLUSIONS

Contemporary signal processors can be used to build sophisticated subminiature speech and hearing aids. Clinical tests proved that the SDSA improves speech fluency significantly. Some methods employing speech spectrum modifications in the auditory feedback loop have been assessed as the most efficient ones. The effects of the device application are noticeable immediately, i.e. speech fluency improves right after the device is turned on. Furthermore, patients are getting more open to interactions with other persons and their social isolation decreases, as well as their school results improve. It should be remembered that most of the patients were school children, so it is of utmost importance that difficulties in learning should be eliminated in the earliest stage of their education.

In addition, the preliminary tests of the special hearing aid pointed out that this device might be effective in hearing loss cases when shifting the speech down the frequency range is the only solution for auditory communication improvement. Otherwise these patients should be provided with a cochlear implant, which is not always possible. Also, in cases of bilateral hearing losses, these patients can be provided both with the cochlear implant and SDSA. However, at this stage, additional tests and experiments are required and are planned.

5. REFERENCES

[1] Andrews G., Garside R., Kay D., The syndrome of stuttering, Clinical Development Clinical, No. 17, pp. 1-191, 1964.

[2] Brennan R., Schneider T., A Flexible Filterbank Structure for Extensive Signal Manipulations in Digital Hearing Aids, IEEE International Symposium on Circuits and Systems, 31 May-3 Jun 1998.

[3] Czyzewski A., Kaczmarek A., Krolikowski R., Skorka P., Automatic Detection of Speech Disorders, Proceedings of the Fourth European Congress on Intelligent Techniques and Soft Computing EUFIT '96, Vol. 1, pp. 183-187, 2-5 September 1996, Aachen, Germany.

[4] Czyzewski A., Kostek B., Roland-Mieszkowski M., Digital Speech Aid Based on the Modification of the Auditory Feedback Loop, Eurospeech' 93, 3rd Eurospeech Congress, 21-23 September 1993, Berlin, Germany.

[5] Czyzewski A., Kostek B., Skarzynski H., IT Applications for the Remote Testing of Hearing (chapter), Springer Verlag, pp. 225 - 247, New York, Heidelberg 2006.

[6] Czyzewski A., Kostek B., Skarzynski H., Internet-Based Automatic Hearing Assessment System, 119th Audio Engineering Society Convention, Preprint No. 6626, 7-10 October 2005, New York, USA.

[7] Czyzewski A., Lorens A., Testing some DSP Algorithms for Hearing Aids and Cochlear Implants, 100th AES Convention, Preprint 4149, 11-14 May 1996, Copenhagen, Denmark.

[8] Czyżewski A., Odya P., Kostek B., New generation aids for laryngectomy patients, 4th Joint Meeting of the Acoustical Society of America, preprint 5aSCa26, 28 November - 2 December 2006, Honolulu, USA.

[9] Czyżewski A., Skarżyński H., Odya P., Kostek B., A Model of Digital Hearing Aid for the Implanted Patients Revealing Residual Acoustic Hearing, 2nd International Conference on Telemedicine & Multimedia Communication, 8-9 October 2004, Kajetany, Poland.

[10] Czyzewski A., Skorka P., Modification of the Auditory Feedback Loop - Effects on Hearing and Speech Production, 100th AES Convention, Preprint No. 4148, 11-14 May 1996, Copenhagen, Denmark.

[11] Donath T.M., Natke U., Kalveram K.T., Effects of frequency-shifted auditory feedback on voice F_0 contours in syllables, Journal of Acoustical Society of America, Vol. 111, No. 1, pp. 357-366, January 2002.

[12] Elman J.L., Effects of frequency-shifted feedback on the pitch of vocal productions, Journal of Acoustical Society of America, Vol. 70, No. 1, pp. 45-50, July 1981.

[13] Harrington J., A Model of Stuttering and the Production of Speech under Delayed Auditory Feedback Conditions, Speech Motor Dynamics in Stuttering, pp. 353-359, Springer Press, New York 1987.

[14] Hartman B.-T., The Neuropsychology of Developmental Stuttering, Whurr Publishers Ltd, London 1994.

[15] Howell P., Changes in voice level caused by several forms of altered auditory feedback in fluent speakers and stutterers, Lang. Speech, Vol. 33, No. 4, pp. 325-338, October-December 1990.

[16] Howell P., El-Yaniv N., Powell D.J., Factors affecting fluency in stutterers when speaking under altered auditory feedback, Speech Motor Dynamics in Stuttering, pp. 361-369, Springer Press, New York 1987.

[17] Howell P., Sackin S., Williams R., Differential effects of frequency shifted feedback between child and adult stutterers, Journal of Fluency Disorders, Vol. 24, pp. 127-136, 1999.

[18] Kaczmarek A., Skorka P., Investigation of the Pitch of the Vocal Tone for Stuttered Speech, 102nd AES Convention, Preprint No. 4486, 22-25 March 1997, Munich, Germany.

[19] Kalinowski J., Armson J., Roland-Mieszkowski M., Stuart A., Gracco V., The effects of alterations in auditory feedback on stuttering frequency. Language and Speech, Vol. 36, pp. 1-16, 1993.

[20] Lee B.S., Effects of Delayed Speech Feedback, Journal of Acoustical Society of America, Vol. 22, No. 6, pp. 824-826, November 1950.

[21] Odya P., Method of Evaluating the Digital Speech Aid Efficacy (in Polish), 10th International Symposium on Sound and Engineering and Tonmaistering, ISSET 2003, 11-13 September 2003, Wroclaw, Poland.

[22] Riper van Ch., The nature of stuttering, Englewood Clifs, New Jersey 1982.

[23] Roland-Mieszkowski M., Czyżewski A., Kostek B., DSA - a new device to decrease or eliminate stuttering, 1st World Congress on Fluency Disorders, pp. 535-539, Munich, 1994.

[24] Soderberg G. A., Delayed Auditory Feedback and Stuttering, Journal of Speech and Hearing Disorders, Vol. 33, No. 33, August 1968.

[25] Starkweather C.W., A simple theory of stuttering, Journal of Fluency Disorders, Vol. 20, No. 2, pp. 91-116, June 1995.

[26] Toccata Plus – Hardware Reference Manual, DSPFactory, 2003.