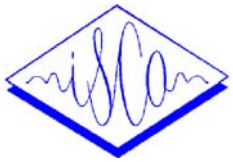

EUROSPEECH '97
5th European Conference on Speech Communication and Technology
Rhodes, Greece, September 22-25, 1997

ACOUSTICAL CHARACTERISTICS OF SPEECH AND VOICE IN SPEECH PATHOLOGY

Natalija Bolfan - Stosic and Mladen Hedjever
Acoustic Laboratory for Speech and Hearing
Department of Logopedics
Faculty of Defectology
University of Zagreb
Kuslanova 59-a, 10000 Zagreb, Croatia
Tel. +385 1 2338 022, FAX: +385 1 229 950, E-mail: mladen@antun.defekt.hr

ABSTRACT

Thirty six hoarse voices of preschool children and fifty speech productions of school children was analyzed using an acoustic analysis by Bruel and Kjaer, Real-time Frequency Analyzer, Type 2123. Thirty six specific oscilograms of sustained vowel productions were divided to oscilograms shapes in three subgroups of specific shimmer values, inside the same group. The purpose of this part of our research is a help in recognition and usage of acoustical terms in diagnostic of disordered voices. Therefore, we obtained "staccato shimmer", the "narrow" total intensity or shimmer, and finally the "wide" shimmer with following oscillations of jitter. The differences in fundamental frequency and intensity between three subgroups of different shimmer were analysed using one-way variance analysis.

The purpose of the second part of this research has been to examine and analyze temporal segments in normal and disordered speech. Temporal segments of school children's speech have been measured starting from the subsound level (VOT - voice onset time and SGD - stop gap duration), and also at the levels of sound, syllable and word. Principal axis analysis showed specific differences between normal and pathological speech in all types of variables.

10.21437/Eurospeech.1997-357

1. INTRODUCTION

The majority of speech disorders manifests itself in specifical changes in fonation and articulation. These changes are most visible in temporal and amplitudinal variations of speech signal. Therefore, the acustical analysis of disordered speech is very usefull in researches, as well as in diagnostics of speech disorders. In speech pathology, the researches of jitter and shimmer that indicates the variations of amplitudinal and temporal oscillations of vocal chords are very useful.

The problem of professionals' ignorance of acoustical terms is often a problem how to make this field closer and understandable. At the Department of logopedics, we use acoustical techniques for speech and voice analyses as a help in diagnostics of speech and voice pathology, as a part of our educational programs, as a main tools in many acoustical researches etc. So, specialists from our acoustical laboratory try to find easiest way to explain acoustical characteristics of speech, voice and different sounds. In this paper, we discuss which are the acoustical indicators of pathological speech and voice, and in which shapes they appear - separately for preschool children's voices and separately for school children's speech.

1.1. Jitter and shimmer

Horii (1979) describes a measure of vocal jitter as a fluctuation in the time interval between the peaks. The same author describes shimmer (1980) as a cycle-to-cycle variation in the amplitudes of the peaks. According to Horii (1979), there is evidence that the magnitude of voice perturbations (jitter and shimmer) in persons with normal voice characteristics is small. Milenković (1987) maintains that it is an intriguing hypothesis that the healthy vocal fold's form produces small periodic oscillations. In contrast, pathological vocal cord produces perturbations of jitter and shimmer. In this paper we discuss about different waveform types. Titze (1993) says that those voice perturbation measures, such as jitter and shimmer, depend on accurate extraction of fundamental frequency (Fo) and amplitude of various waveform types.

The acustical analysis of speech sygnal also can obtain a number of usefull information about some characteristics of speech disorders. It can be supposed that majority of speech disorders also manifests itself in specifical changes of particular temporal segments of speech sygnal. Sometimes, these changes are visible only in "macro-segments"; sometimes, they occure in "micro-segments" of speech and can be measured in miliseconds. They can be shorter than voice itself. These changes are no always significant because of their sound production that sometimes can be perceived; however, their detection and analysing can contribute to comprehensive insight in phenomenology and ethiology of particular disorders, which is the goal of this part of research.

1.2. Temporal and motoric segments of speech

Klatt (1976) analysed and synthetised data from as much as 60 studies dealing with exploration of temporal segments of speech. From his study can be concluded that temporal segments depend of linguistic characteristics od speech expresion, of fonetic context, of individual characteristics of speaker (exp. emotion speech rhythm), and of mothorics of speech mechanism. Crystal and House (1990) concluded that there is no linear correlation between average syllable duration, number of voices in syllable and number of syllables in the word. In last several years, numerous acustical researches of child speech confirmed that various temporal segments in speech tend to shortening with ageing; the variability of these temporal parameters also diminishes with ageing (Smith, 1992; Smith and Kenney, 1994; Smith, Kenney and Hussain, 1996). These changes in childhood are mostly due to neuromuscular maturation and increasing of speech experience (Schwartz, 1995; Smith, Kenney i Hussain, 1996. VOT (voice onset time) also indices the fact that there are a differences between various languages (Flege, 1991;); the characteristical differences occur also at particular speech disorders (Caruso i Burton, 1987; Hedever i Sardelić, 1995; Kent i Rosenbek, 1983) and regarding the age (Paul-Brown i Yeni-Komshian, 1988).

2. METHODS

2.1 Subjects and instrumentary

Measures of fundamental frequency (F0) and intensities (dB) in 36 preschool childrens' voices with voice pathology were obtained using an acoustic analysis by Bruel and Kjaer, Real-time Frequency Analyzer, Type 2123. The subjects were seated in a sound-treated room and their voices were immediately recorded on Real-time Frequency Analyzer. The condenser microphone, Type 4133 (with preamplifier Type 2639) was placed 30 cm from the subject's lips. In this way we obtained 36 oscilograms of group with voice pathology and results of temporal measuring of speech disorders. We have divided these 36 oscilograms of vocal cord's frequency oscillations and intensity in three smaller subgroups as an indicator for diagnostics of disordered voices. These subgroups we called -- the different kinds of jitter and shimmer. In this paper we have presented just 3 voice oscilograms as indicators of each divided group. The speech variables are obtained in syllable, logathoms and word pronouciation, on the sample of 40 school boys aged 10 to 12 with different speech disorders (divided into 4 test groups of ten with dysphonia, dyslalia, stuttering and dysarthria) and 10 boys of same ages, without speech disorders.

2.3 Phonatory Tasks and selections of variables of preschool children voices

There was one phonatory task: 1) Sustained vowel production in which the subjects were asked to articulate vowel /a/ (as long as they can); The vowel was repeated three times with rest periods between vowels. In order the following variables were selected to obtain an acoustical evaluation of the parameters: F0 - fundamental frequency in Hz and DB - intensity of F0 in dB.

2.3. Speech tasks and selections of variables of the school children speech

The duration and intensity of 5 vowels of croation language (A, E, I, O and U), and 2 fricatives (S - like initial sound in english word "sea" and Š like in "shoes") was measured. On CV syllables (combination of stop consonants: P, T, K and vowels) VOT (voice onset time), SGD (stop gap duration) and maximal articulators movement speed in continual fast repeating of syllables.

2.4. Data of voice and speech pattern

The differences in fundamental frequency and intensity in the 3 subgroups with different kind of shimmer were established by a one-way variance analysis. Numerical data obtained from the analyser are submitted to additional statistical analysis (variance analysis, T-test, and multivariate methods: discriminant analysis, regression analysis and principal axis analysis). Because of llimited space, only results of factorial analysis were presented. The both data was processed on PC computer (Program STATISTICA for Windows, Release 4,5 A, (Statsoft, Inc.1993)).

3. RESULTS AND DISCUSSION

3.1 voices of preschool children

According to terminology from the field of voice acoustics, we found specific shimmer values during children fonations inside the same group. So, we obtained "staccato shimmer" (during the interval in which vocal cords do not vibrate - 6 children with voice pathology, figure 1)., the "narrow" total intensity or shimmer (21 children with voice pathology, figure 2.), and finally the "wide" total intensity with following oscillations of jitter (9 children with voice pathology, figure 3.). Pictures of "staccato" shimmer represent very tense fonation (hypercinetic fonation); pictures of "narrow shimmer" represent hoarse and weak voices, and finally the pictures of "wide" shimmer represent loud, excessive, and very hoarse voices. The results from Table 1 show that the 3 subgroups have a statistically significant difference on the following variables: F0 (frequency of F0) and dB (intensity of F0).

Figure 1. Staccato shimmer

A0359G01.gif

Figure 2. Narrow shimmer

A0359G02.gif

Figure 3. Wide shimmer

A0359G03.gif

Table 1. One-way variance analysis of 3 subgroups with different "shapes" of shimmer (level of significance = 0,05) for variables F0 (Hz) and F0 (dB)

Variable	SS Effect	MS Effect	F	P
F0 (Hz)	10352.86	5176.429	9.32434	0.000617
F0 (dB)	906.08	453.042	13.02045	0.000068

3.2. Speech disorders

Because of limited space, only the results of factorial analysis will be presented in this paper. The principal axis analysis by the separate group of variables extracted 8 factors. ***Factor 1*** was obtained by the analysis of variables of vowel A duration (independently of their position in words). ***Factor 2*** *was* extracted from intensity variables of same vowel. ***Factor 3*** represents the duration of remaining 4 vowels (O, U, E and I). ***Factor 4*** was obtained from their intensities. ***Factor 5*** refers to fricative duration, and ***Factor 6*** to their intensity. ***Factor 7*** was extracted from speed variables (fast repeating of CV syllables). ***Factor 8*** represents coarticulation variables (VOT and SGD). After computing the factor scores of 8 factors, we repeated the factorial analysis, and extracted two factors of second level. The eigenvalue of the first factor was 3.442, and the eigenvalue of the second factor was 2.437. The first factor consisted of all three first level temporal factors and speed factor. The second factor consisted of all three first level intensity factors. According to this, factors of second level were dubbed, respectively, as **temporal-speed factor** and **intensity factor.** Three-dimensional figures show great differences between normal and pathological speech (figures 4, 6 and 8) and between speech disorders (figures 5, 7 and 9). Finally, after the factorial analsyis we can conclude that there is a significant differences between normal and pathological speech in at least three dimensions: time, intensity and speed. The specificity of these three dimensions can be showed graphically.

Figure 4.

A0359G04.gif

Figure 5.

A0359G05.gif

Figure 6.

A0359G06.gif

Figure 7.

A0359G07.gif

Figure 8.

A0359G08.gif

Figure 9.

A0359G09.gif

4. CONCLUSION

Thirty six hoarse voices of preschool children and fifty speech productions of school children were analyzed using an acoustic time analysis. The results of acoustical analyses of preschool children voices showed thirty six specific oscilograms of sustained vowel productions. These voice pictures were divided according to following shapes of obtained oscilograms in three subgroups of specific shimmer values during vowel productions, inside the same group. According the purpose of this part of our research, we obtained "staccato shimmer" (tense voice with breakdowns in fonation), the “narrow” total intensity or shimmer (weak and breathy voice), and finally the “wide” total intensity with perturbations of jitter (loud and unpleasant voice). Differences in oscilograms shapes can be useful in diagnostics and therapeutic work with disordered voice. It is important for establishing accurent voice therapy of voice disorders. What is beauty in music, is poverty in the pathology of voice. The differences in fundamental frequency and intensity between three subgroups of different shimmer were established by a one-way variance analysis. The results showed that the 3 subgroups significantly differs on the variables FO (frequency of FO) and DB (intensity of FO).

The factorial analysis of first level factors two second level factors were extracted. The first second level factor was defined by the three temporal and one speed factor of first level. This factor was dubbed **temporal-speed factor.** The second second level factor was defined by the three intensity factors of first level. This factor was dubbed **intensity factor.** The results showed that speech disorders significantly differs in at least three dimensions: time, intensity and speed. It can be conclude that time is one of key dimensions for differentiating the normal and disordered speech. Differences and specificalities of speech disorders can be better defined in complex and dynamic space in interaction of temporal, intensity, and motorical specificities (which include the speed and coarticulation).

REFERENCES

[1] Caruso, A. J., Burton, E. K.: *Temporal acoustic measures of dysrthria associated with amyotrohic lateral sclerosis.* Journal of Speech and Hearing Research, Vol. 30, pp. 80 - 87, 1987.

[2] Crystal, T. H., House, A. S.: *Articulation rate and the duration of syllables and stress groups in connected speech.* Journal of the Acoustical Society of America,Vol. 88, pp. 101-112, 1990.

[3] Flege, J. E.: *Age of learninig affects the authennticity of voice-onset time (VOT) in stop consonants produced in a second language.* The Journal of Acoustical Society of America, Vol. 89, pp. 395 - 411, 1991.

[4] Hedever, M., Sardelić, S. *Okluzija i VOT u govoru kod osoba koje mucaju*. Zbornik referatov: "Logopedija danes za jutri", V. Srečanje logopedov Slovenije, Maribor, str. 85 - 88. 1995.

[5] Horii, Y.: *Fundamental frequency perturbation observed in sustained phonation*. Journal of Speech and Hearing Research, 22, 5-19, 1979.

[6] Horii, Y.: *Vocal shimmer in sustained phonation*. Journal of Speech and Hearing Research, Vol. 23, pp. 202-209, 1980.

[7] Kent, R. D., Rosenbek, J. C.: *Acoustic patterns of apraxia of speech.* Journal of Speech and Hearing Research, Vol. 26, pp. 231 - 249, 1983.

[8] Klatt, D. H.: *Linguistics uses of segmental duration in English: Acoustic and perceptual evidence.* Journal of the Acoustical Society of America, Vol. 59, pp. 1208-1221, 1976.

[9] McReynolds, L. V.: *Functional articulation disorders.* U: Human communication disorders (Ed.: Shames, G. H. i Wiig, E. H.), Charles E. Merill Publishing Company, A Bell & Howell Company, Columbus, Vol. pp. 139 - 182, 1986.

[10] Milenković, P.: *Least mean square measures of voice perturbation*. Journal of Speech and Hearing Research, Vol. 30, pp. 529-538, 1987.

[11] Paul-Brown, D., Yeni-Komshian, G. H.: *Temporal changes in word revision by children and adults.* Journal of Speech and Hearing Research, Vol. 31, pp. 630 - 639, 1988.

[12] Schwartz, R. G.: *Effect of familiarity of word duration in children’s speech: A preliminary investigation.* Journal of Speech and Hearing Research, Vol. 38, pp. 76-84, 1995.

[13] Smith, B. L., Kenney, M. K.: *Assessing duration and variability in speech and non-speech tasks performed by adults.* Journal of the Acoustical Society of America, Vol. 96, pp. 699-705, 1994.

[14] Smith, B. L., Kenney, M. K., Hussain, S.: *A longitudinal investigation of duration and temporal variability in children’s speech production.* Journal of the Acoustical Society of America, Vol. 99, pp. 2344-2349, 1996.

[15] Smith, B.L.: *Relationships between duration and temporal variability in children’s speech.* Journal of the Acoustical Society of America, Vol. 91, pp. 2165-2174, 1992.

[16] Titze, I. R.: *Comparison of Fo extraction methods for high-precision voice perturbation measurements*. Journal of Speech and Hearing Research, Vol. 36, pp. 1120-1133, 1993.