A DEVICE FOR CORRECTION OF RHYTHMICAL DISORDERS OF SPEECH FUNCTIONS

Y.N.GNATIV, Y.M.RASHKEVICH, Z.Y.SHPAK

Department of Automatics
Lvov Polytechnic Institute
Lvov, Ukraine, USSR, 290013

ABSTRACT

The absence of effective approaches to removing human stammering makes it necessary to develop some devices and techniques of raising the efficiency of medical treatment of speech disorders. A device is presented for the implementation of basic approaches used in logopaedics to produce stable rhythmical patterns and rates of speech processes of logoneurosis patients. The performance of the device is based on the introduction of changes into the time and frequency parameters of the speech signal.

INTRODUCTION

The disorders of the speech functions of human beings are considered serious and common diseases.

The accepted view of stammering is as of a stable pathological state of speech [1,2]. Medical treatment of stammering involves a number of procedures, which are designed to destroy stable pathological states of stammering and to create new functional relations, corresponding to healthy speech processes [3]. Here we may include light and sound effects, selection of word pronunciation speed, speech delay, change of qualitative characteristics of speech, etc. According to [2,3,4] an integrated application of these techniques depending on the individual peculiarities of stammering is rather effective.

The above speech treatment procedures require some dedicated technical system. At present the logopaedic treatment makes use of separate devices for time delay of speech, producing periodical sound signals, etc. So far we haven't come across any mention in logopaedic literature of a technical device for increasing or reducing the length of speech pronunciations, in spite of the fact that "slow" speech is considered a classical method of speech treatment [4,5].

The creation of an integrated logopaedic device capable of producing a desired speed of speech reproduction is considered to be the basic requirement in speech therapy.

MAIN FEATURES OF THE DEVICE

Our device is designed to implement following tasks:

- gradual slowing down or speeding up the reproduced speech while retaining its prosodic characteristics;
- introduction of controlled time delays into the output sound signals;
- introduction of additional rhythmical sound or light stimulation of the preset frequency and signal amplitude;
- muffling the speech with "white" noise;
- radical change of the voice quality;
- control of the volume of speech signals.

The device provides two modes of reproducing the speech: variation of the speech rate and time delay. Other corrective modes (rhythm, speech muffling, sound amplification) can be applied independently or in combination with first two ones.

When varying the tempo, the device reproduces a previously recorded text at a higher or lower rate, i.e. increases or reduces the length of sounding without losing the natural quality of the voice, legibility, the key and other speech values. The variation coefficient can be set within the limits of 1-2.5 with the discreteness of 0.1.

If a speech signal is fed into the device in the real time scale, i.e. directly from the speaker, the tempo control produces the displacement of spectral composition of the speech proportional to the variation coefficient. The output speech signal is reproduced at the original speed, but the quality of the voice changes, i.e. an "alien" voice is heard.

The mode of delayed speech enables us to obtain the input speech information at the output of the device with the delay of 100-250 ms (the step being 10 ms). This covers the most favorable range of delays [2] used for the corrective treatment of stammering in patients of varying degree of the disease. An additional amplification of the output sound signals increases the effectiveness of the correcting procedures.

When producing the rhythmical effect the device creates periodical sound and light signals of excitation. The sound rhythms are in the frequency range 0.5-2.5 Hz. During photostimulation light flashes have the length of 70 mc with the rhythm of 2-5 Hz, which is quite suitable for developing necessary correct speech habits [3].

When producing background a casual noise signal is fed into the device output. It prevents the reception of patient's own voice, thus removing pathologically stable reactions to the incorrect speech [2]. To increase the achieved results the patients speech can be recorded and played back at a higher or lower speed for a better evaluation of the deviations.

THE DESIGN OF THE DEVICE

Fig.1 shows the functional layout of the device with the external sound recording and sound reproducing units. When the device is in the process of regulating the speech rate, the input signal is fed into it from the recorder, the latter being controlled by a special circuit according to the preset variation coefficient. Having been converted to a digital code, this signal is recorded by the two memory blocks connected in parallel, each having the capacity of 2K×10 bits. Every new counting is stored in the place of the oldest one. The value of the speech segment in the memory depends on the tempo vaviation coefficient k. The recording frequency is $f_1 = kf_2$, where f_2 - reproduction frequency with a constant value of 16 kHz. The readout from the two memory units is carried out simultaneously, but to different adresses. It

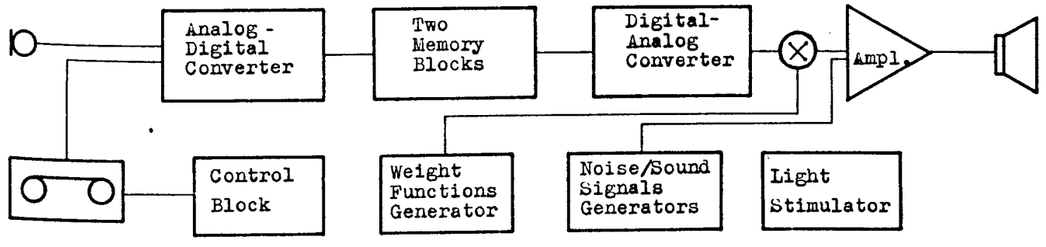

Figure 1. Hardware Description of the Device

enables us to store all the input information in the output signal. The reproduced readings from the memory units are passed to the digital-analogue converter, where they are multiplied by the basic voltage, produced by the weight functions generators. It removes disconnections on the junctions of the speech segments in the output signals [6].

When the device works in the time delay mode, both of the memory units are connected in series. Digital readings of the speech signal received at the input of the device are stored in the memory with the frequency $f_1 = f_2 = 16$ kHz. The difference of the adresses of recording and reproducing is determined by a present time delay. The data obtained from the memory device are fed into the output channels of the system. The generators of sound and noise may be connected to these channels.

THE CONTROL OF THE SPEED OF SPEECH UTTERANCES

According to the phonetic theory of speech formation [7], most of the meaningful information of a speech signal is contained in the transient sections of sounds, the stationary segments being informationally poor. An experiment was conducted to determine the content percentage of various segments in speech utterances. The original speech signal was divided into sections of 20 ms, which equals the length of explosive sounds. On computing the degree of difference between the segments by the DELCO algorhythm [8], these sections were transformed into the output fields, the spectrum composition of which was in fact stationary. The results obtained show that with the normal speech rate the transient sections and short sounds take up about 25 percent of the whole time of the speech signal, the rest being filled with stationary pauses and long sounds. Speeding up or slowing down the tempo in oral speech it take place through the changes of the duration of the stationary sections. The short speech elements change slightly [9]. The idea of regulating the speech rate reproduction makes use of the abundance of speech signals. By excluding short sections of speech signals or by introducing additional short signals it is possible to reduce or increase the time of the phonation of speech utterances, at the same time saving the information content and individual peculiarities as well.

The regulatory process of the reproduction of speech information rate is shown in Fig.2. The original speech signal $x(t)$ is previously stored in some storage device (magnetic tape, digital memory, etc.) during the time of its pronunciation T_p at the rate V_p. The reproduction of the recorded information is carried out during the preset time interval T_r at a suitable rate of $V_r = k V_p$, where $k = T_p/T_r$ - the coefficient of the variation rate. Since it is accompanied by the change of the frequency constituents of the signal, the function of the speech rate regulator is to recover the original spectrum in the output signal $u(t)$ at the time interval T_r.

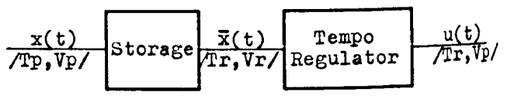

Figure 2. The Speech Rate Regulation Model.

We can single out two groups of methods in the regulation of speech tempo:
- the division of the signal into short uniform sections and changing their length proportionally to the preset rate regulation coefficient;
- the division of the original signal into quasistationary sections with the uniform spectral composition, followed by the variation of phonation time of each of them depending on its original length.

Among the methods of the first group is a selective segmentation of signals [10]. It is simple, easily operated, but has some significant shortcomings. It may remove short sounds and even whole syllables from the speech utterances, which causes distortion of speech and considerably restricts the regulating potential. In combining the segments after the length transformation, phase and amplitude drops occur at the junctions, reducing the quality of the obtained speech.

To remove the possibility of the loss of some useful sections of signals from the speech utterances we suggest using two channel regulation with partial overlapping of the neighbouring segments. To remove the amplitude drops, each of the output sections is multiplied by the weight "window". Partial imposition of segments compensates energy losses in weighing.

The analysis of speech types used for stammering correction shows that all kinds of stammering are characterized by the following regularities [6]:
- the length of every syllable is increased;
- the number of long syllables in a phrase is increased;
- the length of all syllables tends to be equal.

The first two types are positively solved in the presented device. The use of the device permits multiple reproduction of the recorded text with varying speed. On the other hand it allows to listen to an accelerated recording of the patients speech and to determine the intensity of the disorders. Other approaches to speech therapy are possible by varying the time of the sounding of speech information. To make the length of speech syllables equal requires more discriminating approach of the speech signals, which is characteristic of the second group of tempo regulation methods.

CONCLUSION

The performance of the device which permits to carry out a set of correcting logopaedic procedures is described. The most significant of them is the changing of the rate of speech reproduction. The best results in treating speech disoders are to be obtained by using regulator, which makes the length of stationary sections of sounds of equal duration. Modelling confirmed the effectiveness of these approaches allowing to create a wide range of tempo variations while retaining high quality characteristics of the reproduced speech.

REFERENCES

1. Заикание / Под ред. Н.А. Власовой, К.-П. Веккер // М.: Медицина. - 1978. - 200 с.
2. Данилов И.В., Черепанов И.М. Патофизиология логоневрозов // Л.: Медицина. Ленингр. отд. - 1970. - 159 с.
3. Андронова Л.З., Лохов М.И. Использование методов дестабилизации устойчивого патологического состояния в клинике и лечении заикания // Физиология человека. - 1983. - Т. 9. - № 5. - С. 854-859.
4. Куршев В.А. Заикание // М.: Медицина. - 1973. - 159 с.
5. Андронова Л.З., Арутюнян М.А. Анализ временных характеристик видов речи, применяемых при коррекции заикания // Дефектология. - 1984. - № 4. - С. 34-37.
6. А.С. 1173438 СССР. МКИ³ G10L3/02. Устройство для изменения темпа речевой информации / С.В. Балицкий, Т.Н. Гнатив, В.В. Грицык, А.Ю. Луцык, К.М. Рашкевич. - Опубл. 15.08.85. Бюл. № 30.
7. Пирогов А.А. Вокодерная телефония. Методы и проблемы // М.: Связь. - 1984. - 384 с.
8. Маркел Дж., Грей А. Линейное предсказание речи // М.: Связь. - 1980. - 308 с.
9. Агафонова Л.С. и др. О некоторых характеристиках русской речи в зависимости от темпа произношения // Слух и речь в норме и патологии. Л.: Наука. - 1977. - С. 25-39.
10. Шифман М. Регулятор темпа речи воспроизводящего магнитофона // Электроника. - 1974. - № 17. - С. 24-35.