



Computer programme for speech impediment diagnosis and therapy

Mariusz Dzieńkowski^{a*}, Wiesława Kuniszyk-Jóźkowiak^b,
Elżbieta Smółka^b, Waldemar Suszyński^b

^a*Management Department, Technical University of Lublin,
ul. Nadbystrzycka 38, 20 – 618 Lublin, Poland*

^b*Institute of Physics, University of Maria Curie-Skłodowska
Pl. Marii Curie-Skłodowskiej 1, 20 – 031 Lublin, Poland*

Abstract

The article contains a description of a computer programme which supports diagnosis and therapy of stuttering. The application of computer technology allows for carrying out current analysis of the speech fluency disturbance level in the course of the whole therapeutic procedure. The main features of the process are: identification, classification, record of non-fluent fragments in the form of sound files and creation of a database of fluent counterparts and instructions. Most of the tasks are performed by a computer, controlled by the programme presented herein.

1. Introduction

Diagnosis and therapy of stuttering are difficult tasks to perform because of the great complexity of the disturbance. It results from the fact that the examined people are characterised by different creation of speech structures as well as character and kind of speech process disturbances. Many patients, especially adults, use supportive procedures, which consist in avoidance of certain words and sound combinations or sentence shortening. Apart from that, stuttering depends on many external factors and on the patient's mental and physical conditions [1]. The intensity and kind of the occurring disfluencies may vary depending on the day and situation. For these reasons, logopaedists do not carry out therapy willingly.

The presented herein "Computer programme for speech impediment diagnosis and therapy" is a solution to these problems. It has been created to facilitate current analysis of the speech fluency disturbance level and to create a possibility to monitor constantly and individualise the therapeutic process. Therapy carried out with the use of this tool consists of recording a stuttering

* Corresponding author: *e-mail address*: zeb@antenor.pol.lublin.pl

person's utterance (e.g. reading a text, telling a story, answer to a question) and automatic generation by a computer (display and utterance) of tasks to be performed by the patient. The assessment of the speech disturbance level (i.e. diagnosis), on the other hand, which is performed by a logopaedist, supported by the programme, consists of the following stage identification of particular non-fluent words, their record as sound files for further analysis and classification, and register of fluent counterparts, used in designing instructions for therapy. For the fact that stuttering should not be considered solely as speech fluency disturbance, tasks improving the speaking process and generating fluent speech should be linked with psychological and pedagogical treatment etc. Such elements may be included in the program by the logopaedist, by means of appropriate control task and instruction design.

2. Description of the programme for speech impediment diagnosis and therapy

The computer programme for speech impediment diagnosis and therapy is a multi-window system working in the Windows environment. It is easy to operate and has a friendly interface. It has been created with the Delphi 5.0 packet – a RAD (Rapid Application Development) type visual tool [2]. The programme was coded in the Object Pascal language. Moreover, the *Playback* and *Recording* modules employ API functions and procedures, which serve voice communication between the system and a sound device.

Launching the *Terapia.exe* (*Therapy.exe*) file causes the programme to start and show a splash form of the application for a few seconds, containing the title and names of the authors. Then the main window shows with buttons which enable the user to access the three main modules of the programme, i.e.: therapy, diagnosis and instruction. At the first launching of the application, a directory structure is created, where data of the programme will be stored as sound and text files. Three subdirectories are created in the main catalogue: *Instructions*, *Fluency* and *Subjects*. In the first one, recordings of diagnostic tasks will be stored, in the second – fluent counterparts of the non-fluent fragments of utterances, and in the last one – subdirectories with patients' names. Each examined person will have his or her own directory. A sound file will be stored there, with constantly updated information on the course of the patient's therapy as well as subdirectories with the dates of therapy, containing a registered utterance and its non-fluent fragments. Recorded utterances, non-fluent words and their fluent counterparts will be saved on the disk as WAVE files. The optimum parameters for registering speech in the memory unit of the computer are the following: sampling frequency – 22050 Hz, resolution – 16 bits, number of channels – 1 (mono). Information on the sound parameters and directory location is stored in the registry of the operating system.

3. Sound recording and playback (sound communication human – computer)

Sound recording and playback have been realised in two distinct units, so called Thread Objects. Thanks to this approach they are independent threads and they do not block the system when it plays large sound files back. Moreover, in order to avoid time-consuming disk I/O operations, the so-called mechanism of memory mapped files was applied. This feature of Win32 allows the user to refer to disk files in the way as if they constituted a fragment of the random-access memory. Sound recording and playback in real-time is implemented by means of double buffering. To do this, at least two memory buffers are necessary. This method allows for simultaneous recording of data into one buffer and processing by the other during the record. Playback also needs the circular buffer method because of the great size of wave files, which should not be loaded in memory as the whole. A complete set of functions for processing audio signals for Windows operating systems is provided by Microsoft. The sound recording process consists of a few stages:

- determine the sound recording format by filling the `WaveFormatEx` wave structure,
- open the `WaveIn` device and get handle to this device (`WaveInOpen()` function)
- define the special header structures, which determine, among other things, the location and size of memory buffers and their readiness for use in the system (`WaveInPrepareHeader` function),
- add headers to the buffer system (`WaveInAddBuffer` function),
- start recording (`WaveInStart` function), buffers record into the file,
- stop and finish recording (`WaveInReset` and `WaveInStop` functions),
- free all used memory and unprepare the buffer headers (`WaveInUnPrepareHeader` function).

Playback, which is an easier process, is realised by means of the following functions: `WaveOutPrepareHeader`, `WaveOutWrite`, `WaveOutPause`, and `WaveOutRestart`.

4. Description of the programme operation and handling

4.1. Therapy

The therapy process begins with the display of the *Subject* window, where we mark or find the person on the list (Figure 1). The buttons which we find on that form implement the following functions:

- *Description* – display of information contained in the text file concerning the subject and course of therapy,

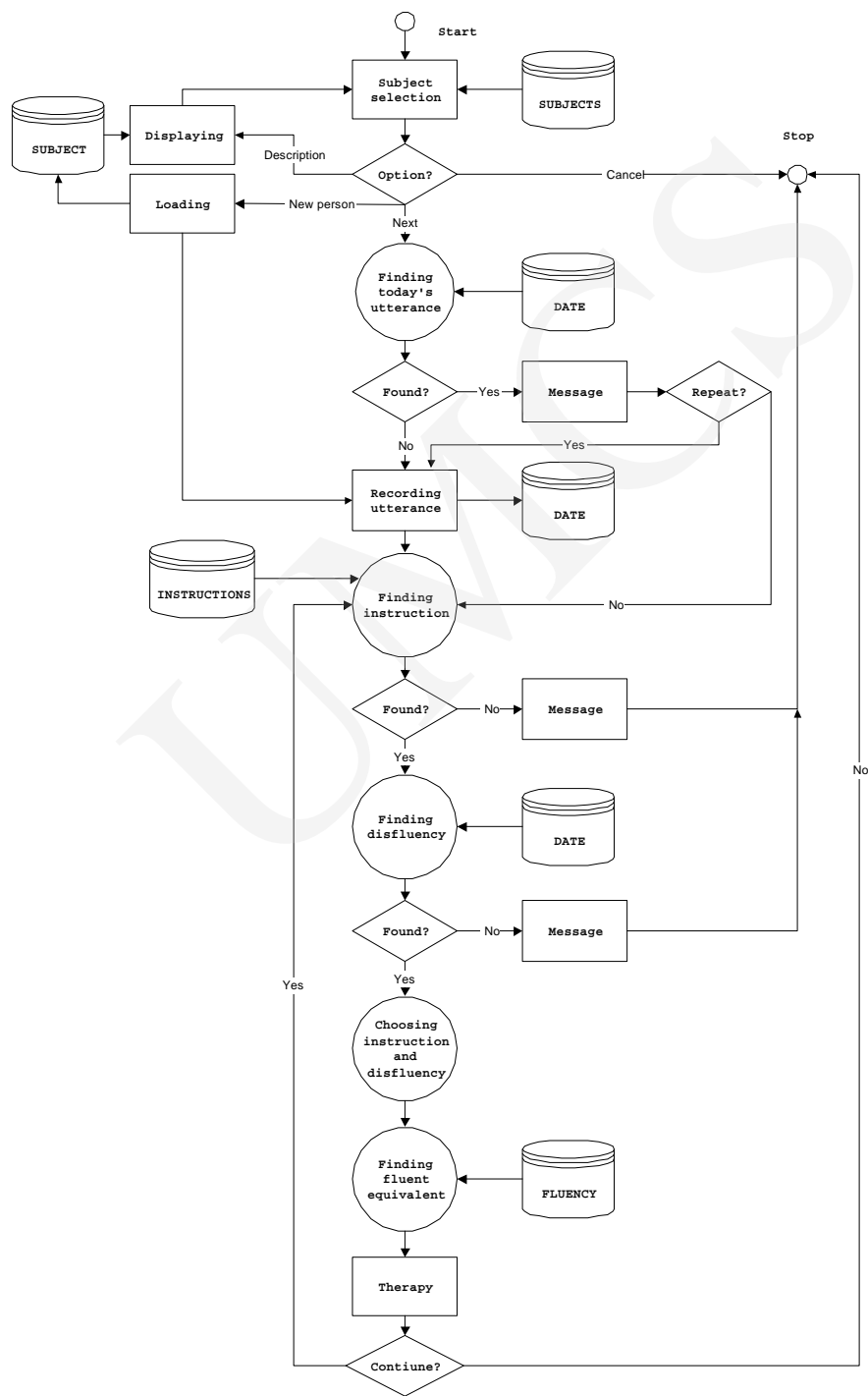


Fig. 1. Therapy module diagram

- *Next* – continuation of therapy, display of the window implementing the utterance record,
- *New subject* – display of the form to fill in with the data of a new patient, creation of a directory and text file with therapy description,
- *Cancel* – end of therapy.

If a person is not selected, the *Next* and *Description* buttons will not be active. After the subject is chosen or a new one is added, in his/her directory a subdirectory will be created with the date of the therapy session. Then the programme will display the *Utterance record* window, where the task will be implemented. In this form we have two buttons to choose from: *begin* and *finish record* as well as a field, where the title or kind of utterance should be typed in. When the record has been saved, the therapy as such begins. A window is displayed on the screen, where tasks for the patient appear. Apart from that, the tasks are also at the same time uttered by the computer. The diagnostic tasks are generated as follows:

- an instruction is randomly chosen from the *Instructions* directory,
- a non-fluent word is randomly chosen from the directory with the previous therapy date and then its fluent counterpart is found in the *Fluency* directory,
- instruction linked with the fluent word is displayed and played back.

If the computer fails to find a disfluency or instruction, an appropriate message is displayed. Continuation or finishing the therapy session depends on the choice of the *Continue* or *Finish* button. The search for particular directories is done by means of the following functions: `FindFirst()`, `FindNext()` oraz `FindClose()`.

4.2. Diagnosis

Choosing the *diagnosis* button causes the programme to display the *Subject* window in the main window with the list of people, list of dates of therapy sessions and 3 buttons: *Next*, *Description* and *Cancel* (Figures 2 and 3). Choosing an item in the main list activates the *Description* button and allows for display and, optionally, modification of the information concerning the patient and his/her therapy. The choice of a date from the second list and pressing the *Next* button will allow the user to continue the diagnosis. Its next stage is the graphical presentation of the sound file with the utterance in the form of an oscillogram. When the window appears, new elements are added to the tool bar, i.e. *play*, *pause*, *stop*, *enlarge* or *decrease scale* buttons and edit fields which determine the borders of the marked fragments of the recording and the scale.

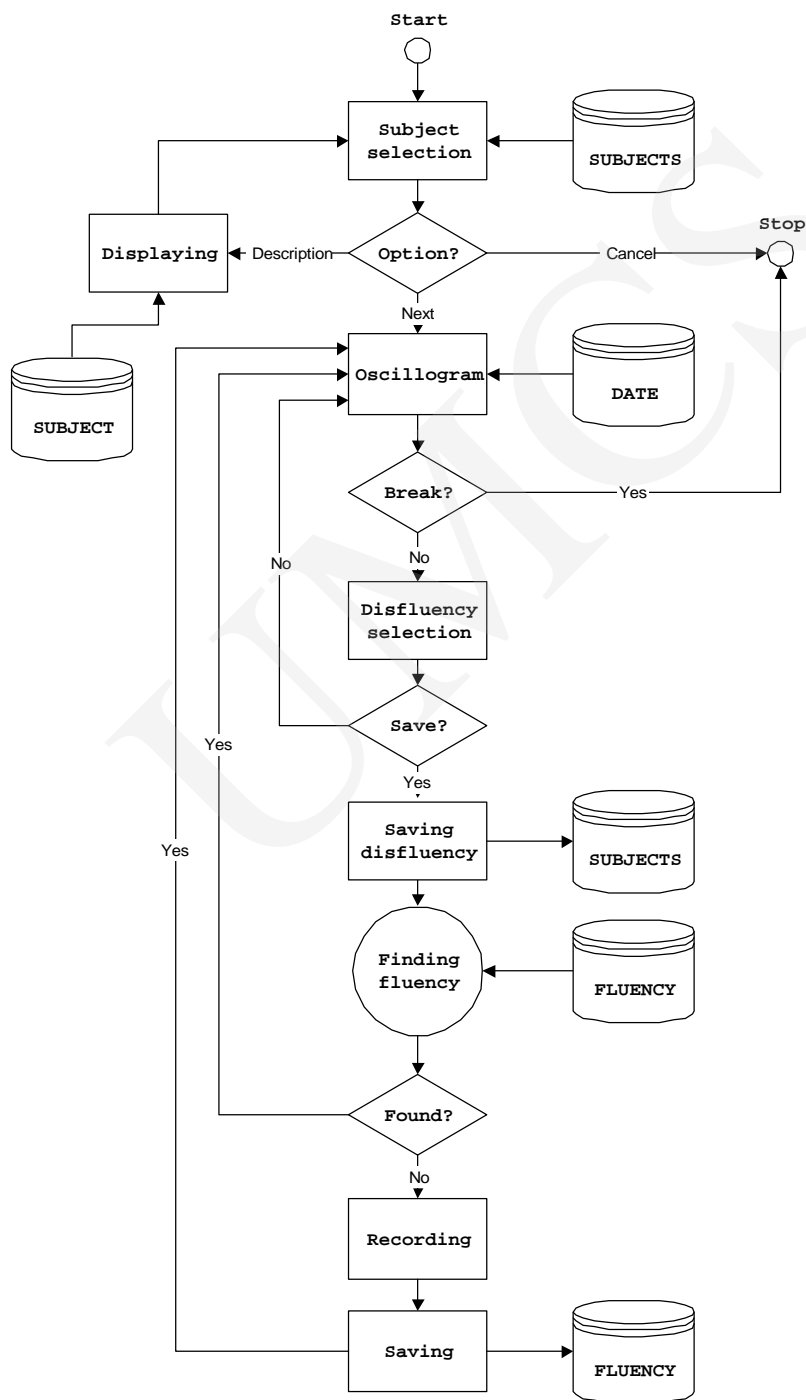


Fig. 2. Diagnosis module diagram

The oscillogram window is divided into two parts: the lower part shows the whole content of the sound file and the upper one allows for scaling and scrolling the graph. Below them a status bar has been placed, containing information on the file size and format. The graphics of the form is implemented with two components of PaintBox. The advantage of the objects is their speed, and the disadvantage is the need for constant refresh of the whole image. Maximum speed of oscillogram creation has been achieved thanks to the applied mechanism of file mirroring in the computer memory and the graphic method of double buffering.

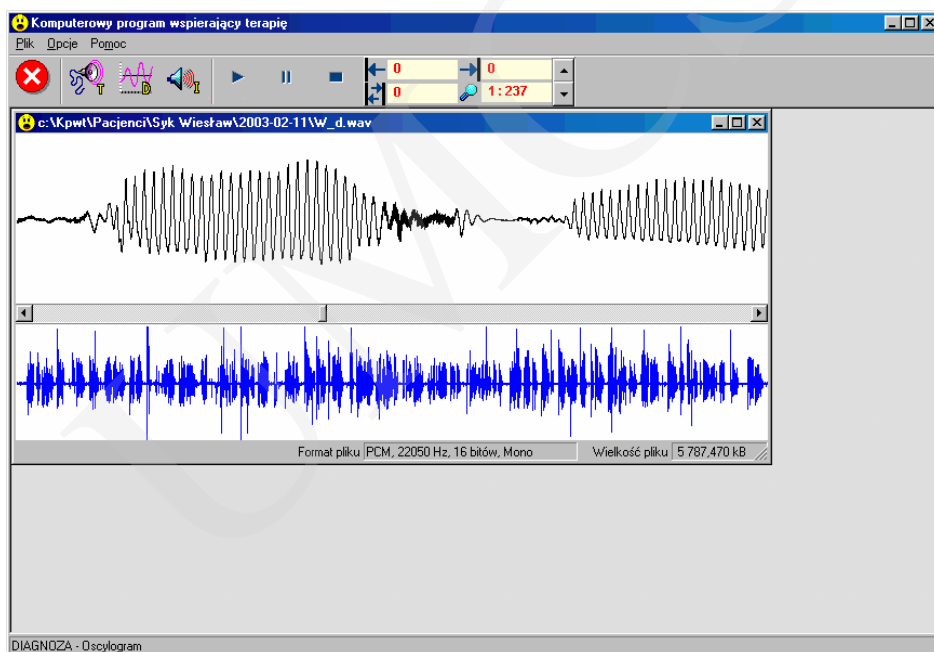


Fig. 3. *Diagnosis* module window

This graphical way of presentation of a stuttering person's utterance which is the easiest allows the logopaedist to choose the non-fluent fragments and save them. The borders of the non-fluent words are marked with the mouse. Right-clicking the mouse in the marked area will activate a popup menu with the following options: *Cancel*, *Play*, *Save disfluency*. When the last option is chosen, a *Save non-fluent word* window is displayed, where a fluent counterpart should be entered. Then a fluent word which corresponds to the saved disfluency is found. If the search result is positive, the programme goes back to the oscillogram. If the programme fails to find a fluent counterpart, a fluent word is recorded. The diagnosis is finished when the graphic window is closed.

4.3. Instructions

The third module of the programme serves creation, deletion and audio editing of therapeutic tasks. These tasks are implemented in the *Instructions* window (Figure 4).

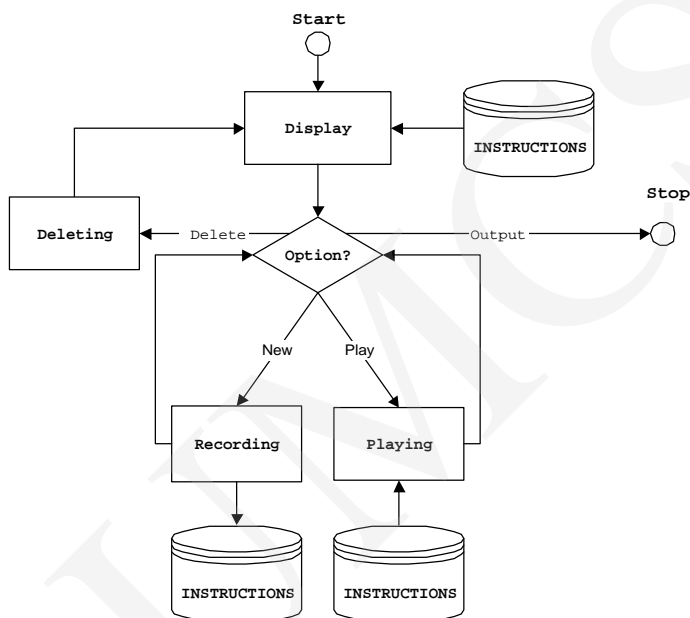


Fig. 4. *Instructions* module diagram

5. Conclusions

In the method of stuttering diagnosis and therapy elaborated by the authors of the work, many tasks are performed by the computer. It aims at making it easier for logopaedists to perform the difficult operations. The presented method may also be applied in the correction of other speech impediments. The authors plan to develop the programme further by adding a correction factor such as the echo, speaking in chorus with the computer, repetition etc [3]. Moreover, the authors are working on automatic detection of particular disfluency kinds in continuous speech, which will greatly facilitate diagnosis [4].

Acknowledgements

The research was supported by Grant No. 4 T11E 035 22 from the State Committee for Scientific Research in Poland.

The authors thank Natalia Fedan M.A. for preparing the text in English.

References

- [1] Bloodstein O., *A Handbook on Stuttering*, Singular Publishing Group, Inc., San Diego, California, (1995).
- [2] Osier D., Grobman S., Batson S., *Delphi 2*, Helion, Gliwice, (1996).
- [3] Adamczyk B., Kuniszyk-Józkowiak W., Smółka E., *Influence of echo and reverberation on the speech process*, Folia Phoniatr., 31 (1979) 70.
- [4] Suszyński W., Kuniszyk-Józkowiak W., Smółka E., *Acoustic methods in diagnosing and therapy of speech disorders*, Maintenance and Reliability, 7 (2000) 19.