The speech prosody of people with stuttering and developmental apraxia: the efficacy of an intervention program

Bernadette Cardoso and César Reis

Universidade Federal de Minas Gerais
Phonetic Laboratory
bernavon@ciclope.lcc.ufmg.br / creisufmg@gmail.com

Abstract

This study has the aim of developing knowledge, methodology and technique in the prosody area, for its application on the speech therapy, with the assistance of a feedback system – with *WinpitchPro* (Philippe Martin). A database was constituted on the basis of a patterned *corpus* (EUROM1) by its translation, recording and editing on .wav format. Throughout model-sentences, an experimental intervention program initiates with the participation of 2 adults with stuttering problem and 2 with apraxia of speech. The efficacy of the therapeutic procedure is tested on the basis of the acoustical and the statistical analyses which have compared the first and the second speech samples taken before and after the twenty (20) therapy sessions held in a prosody approach only.

Index terms: Speech, prosody, intervention, apraxia, stuttering

1. Introduction

The professionals of speech and language disorders always face challenges of explaining the causes of speech problems as well as searching ways of reducing or even solving the effects they have over one's life. Stuttering and apraxia of speech are among others, language impairments which the origins are not well known and the solution uncertain. Different intervention programs have been proposed and, although their results are positive, they are not comprehensive at all. This study has the aim of presenting a therapeutic procedure based on prosodic resources to which adults who stutter and adults with apraxia of speech can participate. After the program is proposed and applied, the outcomes are examined according to its therapeutic efficacy.

The program was designed to develop knowledge, methodology and technique in the prosody area for its application on the speech therapy with the assistance of a feedback system of the prosodic patterns. We assume that prosody is the structuring element of verbal production and its faculty of connecting speech sounds and establishing a hierarchy between syllables, that ends up generating a rhythmic structure which can facilitate even the short term auditory memory.

Furthermore, if in fact prosody facilitates speech production, we also assume that once the individual with communication problems becomes acknowledged to the prosodic organization in its descriptive units, he becomes capable of focusing in the prosodic aspects and, in advance, of improving his verbal performance.

If on one hand, we take the prosody as a cohesion element of speech sounds [1] [2] syllables, words and phrases, on the other hand the transient aspect of the sound, especially concerning the melodic traces of speech, precludes its therapeutic use. Once a melodic contour is produced, it becomes hard to recover and talk about it. Psycholinguists have long concluded that not only the sounding aspects, but also words and their syntactic ordination, are out of memory seconds after they are thrown out on the air. After an enunciation, what remains in one's mind is just the idea or the origin of that communicative intent [3].

Taking these facts into account, we adopted the *WinpitchPro* [4] - an applicative which does the prosodic analyses in real time – in order to document the speech event. Once we get the visual image of Fundamental Frequency (F0) curve on the computer screen and the recordings, it becomes quite feasible for the individual to comprehend and to do the verbal changes required from him.

2. Methodology

In order to do this research, we constituted a database by translating, recording and editing the passages of the EUROM1 *Corpus* [5]. Two speakers – a male and a female - have recorded the material and we edited it on the *Sound Forge* 8.0 program. Two experimental groups were also constituted – two (2) adults with stuttering problem and two (2) with apraxia of speech, without any other disorder, participated on the study. We selected these two speech language disorders because of the existing similarities amongst them - the incidence by gender, the speech patterns they show having both segmental and suprasegmental problems [6] - and also because we assume their problems have the same nature and can be attributed to the same hierarchical level of motor control for speech [7].

Each one of the participants had their speech sample collected, individually, in a sound proof cabin. For the recordings we used a DAT (Digital audio tape) TDC-D7 model and a microphone ECM/Sony. After the first sample was collected, these subjects participated on twenty (20) therapy sessions for a three month period. At the end of the intervention program, the second speech sample was collected. Over these two samples, both having reading and conversational speech, an acoustical analysis was done as well as the statistical analysis for the comparison of the results. The applied model was the single subject design, where the individual is his own control, taking into account his performance in pre-treatment condition [8] [9].

During therapy sessions, which took place in an individual manner, the participant listened to an integral EUROM1 passage. Afterwards, he was presented to its constitutive period on the screen which was a sentence model he had to

reproduce. Every utterance was reproduced and recorded, without any limit to the number of tries.

In order to lead the participant to produce an utterance equivalent to the model, the therapist used several of the resources offered by *Winpitch* as follows: a selection of a block on the screen; the repetition of the selected portion as needed; the reproduction in MIDI mode; the reproduction in SLOW mode and the presentation of the two windows simultaneously for the comparison between the model and the attempt of reproducing it.

The judgment, about the satisfactory quality of the model reproduction, was made on a consensus basis between the therapist and the subject. In order to do that, they observe the direction and the extension of the melodic movement (F0 reached) as well as the reproduction of the same stressed syllables, keeping the same rhythmic pattern of a given model. Along the task of reproducing the model, some guidance was followed. We have noticed that the task of identifying the number and the specific salient syllables of the model would facilitate not only the reproduction but also the recovery of the sentences especially the long ones. Because of that, the first attempt of reproducing the model was always proceeded by this auditory identification. Others instructions were also given as the following:

- Give preference to speech rhythm rather than articulatory precision: every time the subject did not achieve the model, following the same process applied by the modeling utterance, the therapist asked him to make another try. Following this instruction, the participant became aware of the application of the phonological processes is proper to the connected speech and he shall use them instead of focusing on articulatory precision of speech segments they are used to.
- Focusing on the ongoing production: while the subject tries to reproduce the model, the therapist asked him to always take into account the kinesthetic cues of speech rather than anticipating production problems, which is always the focus of these populations.
- Given attention to the final portion of the utterance, where the word which contains the more important content meaning is located and to which belongs the syllable with greater intensity, greater duration and the more important melodic movement. This advice has the intent of making an opposition to these individual's difficulty of initiating an utterance. Individuals who stutter as well as those with developmental apraxia diagnosis very frequently report they fear the beginning of an utterance. Although this difficulty is not well established, it might be related to their inability to make the physiological changes necessary for speech production from a resting position. We assume that being aware of the importance of the final portion of the tonal group would attenuate the stress relative to the beginning of an utterance and it would promote the accomplishment of a more fluent speech.
- Finally, the subject was guided to never stop a production before its end. In this sense, the speech is considered a performance and being so, it can not be interrupted before it is completed.

For the statistical analysis of the reading and conversational speech performed over the first and the last samples some measures were taken. The durational measures include the syllable duration - stressed and non-stressed syllables, the duration of the disfluencies, the pauses duration and the total duration of the utterances. The fundamental frequency measures took into account the initial, maximal and minimal frequency of the utterance as well as the entire melodic movement of the prominence. All the measures were taken manually by the first author on the *WinpitchPro* program.

All the data were noted down on a paper spreadsheet and typed on Excel lately. For the statistical procedures, we decided to consider the shorter instances of disfluency, which always tend to be the last token of a series of repetitions. The articulation rate measure, which we consider the best estimate of time in speech execution, was calculated by the total time of an utterance divided up by the total number of syllables after pauses and instances of disfluency have been taken away. We also have considered the phonological syllables so that the vocalic sandhis (pouco espaço = /ˈpoʊ̯kɪsˈpasʊ/) have resulted in higher speech rate.

The acoustical analysis was done through the *WinpitchPro* and the statistical analysis was done on the *Mini TAB* and the *R program.*

3. Results

The results have shown that the therapeutic program has contributed for some modifications on the participants verbal performance.

In regarding to durational parameters, the analysis performed over the speech sample pre and posttreatment revealed that an apraxia subject (NS) has increased significantly his stressed and unstressed syllable duration (p-value= 0,0076). Also, a subject with the more severe stuttering problem (FD) has decreased the unstressed syllable duration. For this individual only the duration of stressed syllables has decreased at statically significant level (p-value= 0,0002).

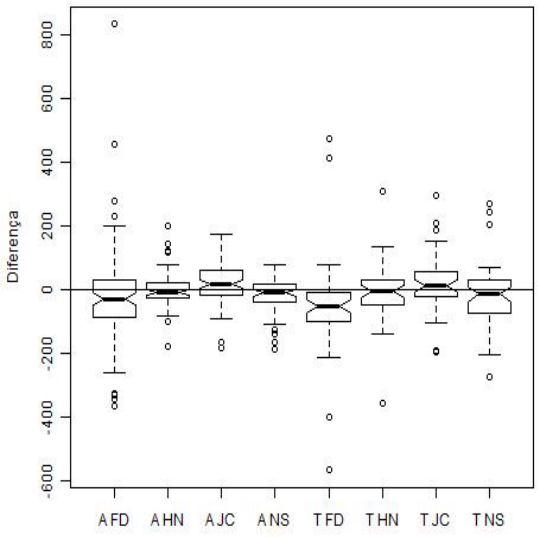

Figure 1. *Mean comparison of the difference between pre and post treatment: stressed and unstressed syllable duration (ms) by individual*
FD e HN (stuttering group) - JC e NS (apraxia group)

The comparison of the differences pre and post treatment conditions on the graphic 1 illustrates the changes in regard to syllable duration for each one of subjects. HN, a subject with a mild stuttering problem, has showed no changing as long as all his data is located next to zero level. On the other hand, the modification put into effect by FD is remarkable, especially concerning the duration of the stressed syllables, since over than 75% of his data fell below zero level of reference after treatment.

The change on the durational parameter exhibited by a subject with stuttering (FD) and by another with apraxia (NS) was also demonstrated on the comparison of the speech rate and the articulatory rate pre and post treatment (graphic 2). In

this variable, as well as in others, there seems to be a tendency followed by the group, although only those subjects have showed statistically significance in modifying their speech rate.

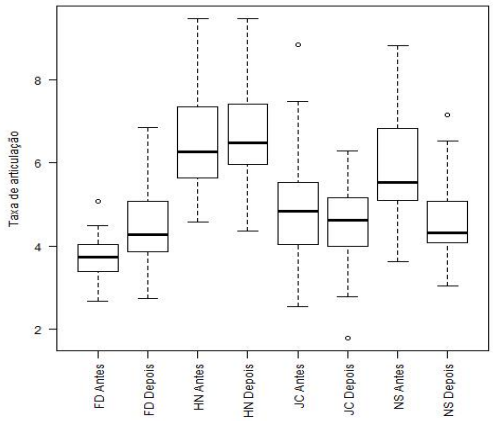

Figure 2. *Individual data; pre (A) and post treatment (D) condition on Articulatory rate (syllables/milliseconds)*

The analysis of the disfluencies was held only for the stuttering group because people with apraxia have not produced sufficient difluencies for the application of statistical procedures.

Taking into account the absolute number of disfluencies, we observe a decrease in the frequency of stuttering after treatment. If we examine duration of the speech disfluencies, we also find that the instances become shorter after intervention.

Table 1. *Mean comparison of disfluencies duration (in milliseconds) pre and post treatment condition for the stuttering group*

Treatment	Mean (Standard deviation)			
pre /post	N	Mean	StDev	p-value
pre	127	819	733	0,001
post	40	493	427	

Although the participants with apraxia have exhibited unexpected changes on the durational variables, they have shown a behavior consistent with progress on the F0 parameter on the reading condition. For both subjects with apraxia, the melodic movement of the prominence revealed greater extension at statistically significant level.

Table 2. *Mean comparison of the Extension of melodic movement (in Hertz) of the nuclear syllable pre and post treatment for JC subject (Apraxia group)*

Treatment	Mean(Standard deviation)			
pre / post	N	Mean	StDev	p-value
EMMP/pre	7	13,57	4,93	
EMMP/post	7	32,00	5,24	0,010
Difference	7	-18,43	13,10	

Table 3. *Mean comparison of the Extension of melodic movement (in Hertz) of the nuclear syllable pre and post treatment for NS subject (Apraxia group)*

Treatment	Mean(Standard deviation)			
pre / post	N	Mean	StDev	p-value
EMMP/pre	7	62,1	35,6	
EMMP/post	7	84,3	37,3	0,001
Difference	7	-22,14	10,01	

Likewise, they have elevated the initial frequency of the sentences they produced after treatment. They have also extended the overall pitch range of their utterances at significant level in reading condition. In this aspect they were accompanied by the subject with the mild stuttering problem (HN), who extended his tessitura from 84.7 Hz before to 118.4 Hz after treatment.

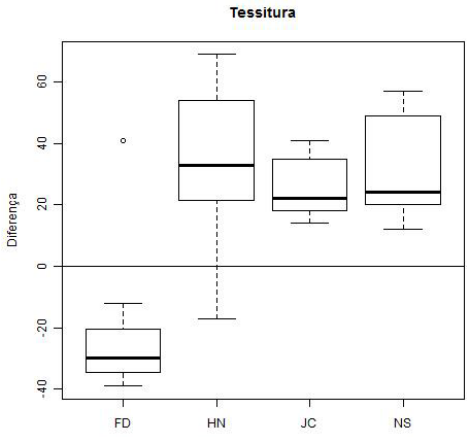

Figure 3 – *Comparison of differences on the overall pitch range of the utterances. Results by individuals on speech reading samples*

On the comparison of differences, the plotted data on graphic 3 revealed clearly that only FD, the subject with the most severe stuttering problem and who have showed the greater changes in the durational parameter, to the contrary he has reduced the overall pitch range of their utterances after intervention.

4. Discussion

The comparison of pre and post treatment data indicates an improvement on the disfluency problem for these subjects with stuttering. The greater progress was accomplished by FD, the individual with the most severe problem. Especially because he has reduced the frequency of *inaudible sound prolongations* and *sound syllable repetitions*, and the duration of *audible sound prolongations*.

The improvement in the speech disfluencies of those subjects with stuttering happened not only because they have diminished the frequency but also because they have reduced the duration of their instances of disfluency. This means that, although they were still stuttering after treatment, this happens 55% less than in the beginning and their instances became shorter.

In addition, the changes exhibited on the durational parameter by those with stuttering problem were followed by a significant reduction on the duration of stressed and unstressed syllables which results on an increment of speech velocity (speech rate and articulatory rate). Although the other participant with stuttering has showed the same tendency, the exam of individual data revealed that FD data has strongly contributed for reaching the statistical significance.

The progress on the temporal aspects of speech reached by the individuals with stuttering can be considered more comprehensive because, for those, the applied program promoted changes in conversational speech and reading speech as well.

The modifications on the Fundamental frequency (F0) parameter, implemented by 75% of the participants, were not transferred for the conversational condition. Notwithstanding, in the reading condition, three out of the four participants have

significantly increased their tessitura, this larger overall pitch range did not take place for any group in spontaneous speech.

The results, as they have showed, make us to conclude that the modifications on F0 were task-specific, without any generalization from speech reading to conversational speech. By the time we were collecting the second speech sample, two of the participants mentioned that they have considered the three months of treatment a very short period of time for the experimental intervention. Probably, if the interventional program had lasted for a longer period of time, it could result in a broader speech range use in spontaneous speech as well. However, this explanation has only a speculative nature until the application of the program for a longer period is tried.

5. Conclusion

On the 70's, several therapeutic program had been proposed under the same approach denominated "fluency therapy". In that kind of intervention, the individual who stutters was encouraged to promote a reduction of his overall speech velocity deliberately. As a result of this procedure, a drastic reduction of the stuttering instances was obtained. Lately the use of electronic devices such as *Speech Easy* has led to surveys of use satisfaction as well as to efficacy studies concerning the reduction of stuttering frequency [10]. Although the goal of fluency can be reached throughout these procedures, the awkwardness and the lack of speech naturalness make people abandon these techniques completely.

In our work, we have given preference to the prosodic aspects of speech, which also have to do with speech velocity. However, the changes we have proposed, which result from a greater control of speech themselves do not disturb the verbal naturalness and verbal expressivity. The attention given to the melodic movements and to the rhythm inherent to a normal speaker must guarantee a naturally expressive production.

Our program intervention, in the manner it was suggested, is based on the neurophysiologic definition of stuttering on regard to its argument of an existing dysfunction of speech automatism [11]. If in fact there is a deficit on the automatic control of speech for those people who stutter and those with apraxia; and, as a result of that, their verbal production must happen under a voluntary control; it would be very beneficial for them to focus on the prosodic parameters of speech. By these means the naturalness of language is preserved with its adequate rhythm and melodic contours.

A preliminary question we had to undertake this work, which was partially answered, has to do with the modeling in language acquisition. Studies concerning first language acquisition as well as developed methods for teaching a second language make us aware of the importance of modeling. In his interaction with a young child, the adult models ahead of time utterances that the child only will be capable to produce later on [12].

Likewise, teachers of second language lead their students throughout structured models which become functionally acquired in a very gradual manner. As long as we have taken utterances as those of EUROM1 *corpus*, which have an approximate size of a paragraph and which are meant to convey a daily conversational speech; and, as we have prepared the recordings of the models along with actors experienced with the expressivity of verbal material; it seems that, from the reached results, the proposed program is very promising in its task of suiting the production of adults either in a second language or on the disordered speech.

6. Acknowledgements

This work was supported by the FAPEMIG (Fundação de Amparo à pesquisa do Estado de Minas Gerais), Grant # SHA 0097/06.

7. References

[1] CONTURE, Edward. Stuttering: Its Nature, Diagnosis & Treatment. Needham Heights, MA: Allyn & Bacon. 2001. p. 444.

[2] HALLIDAY, Michael A.K; HASAN, Ruqaiya. Cohesion in English. London:Logman, 1976.

[3] FODOR, J.A, BEVER, T.G. The psychological reality of linguistic segments. Journal of verbal learning and verbal behavior. v. 4, p. 414-420, 1965.

[4] MARTIN, Philippe. WinpitchPro (Version 2006).

[5] EUROM.1 database. European ESPRIT- Project 2589 (SAM), 1993. Translation held by the staff of the *Laboratório de Fonética* of the Faculdade de Letras da Universidade Federal de Minas Gerais.

[6] CARDOSO, Bernadette von Atzingen S; REIS, César. Variables for the study of the temporal organization in speech disorders. In: Speech Prosody 2008. Anais Fourth Conference on Speech Prosody, Campinas-Brazil, 2008.

[7] SHRIBERG, L., ARAM, D., KWIATKOWSKI, J. Developmental Apraxia of Speech: II. *Journal of Speech and Hearing Research*. v. 40, p. 286-312, 1997(b).

[8] HEFFNER, Christopher L. Research Methods for Education and Social Sciences. AllPsych Online, 2004.

[9] JANOSKY, Janine E., LIBKUMAN, Terry M., LEININGER, Shelley L., HOERGER, Michael P.. Single subject Designing in Biomedicine. Springer Science. 2009, p. 124.

[10] ARMSON, Joy., KIEFTE, Michael. The effect of Speech Easy on stuttering frequency, speech rate and speech naturalness. Journal of Fluency Disorders, v. 33 (2), p. 120-134, 2008.

[11] ALM, Per A. Stuttering and basal ganglia circuits. Journal of Communication Disorders. v. 37, p. 325-369, 2004.

[12] NELSON, Keith,. CAMARATA Stephen M., WELSH, Janet., BUTKOVSKY, Laura. Effects of Imitative and Conversational Recasting Treatment on the Acquisition of Grammar in Children With Specific Language Impairment and Younger Language-Normal Children. Journal of Speech and Hearing Research. v. 39. p. 850-859, 1996