

Automated assessment and treatment of speech rate and intonation in dysarthria

Heidi Martens, Gwen Van Nuffelen, Marc De Bodt
Rehabilitation Center for Communication Disorders
Antwerp University Hospital
Edegem, Belgium
heidi.martens@edpnet.be

Tomas Dekens, Lukas Latacz, Werner Verhelst
Department of Electronics and Informatics
Vrije Universiteit Brussel
Elsene, Belgium

Abstract—Prosody assessment and treatment in dysarthria is clinically relevant, since prosodic impairment can have a negative impact on speech intelligibility and thus on participation in daily life conversation. We propose a speech-technology based software tool that provides automated numerical and visual feedback on two important aspects of prosody: speech rate and intonation. The tool includes speech rate and intonation algorithms, both specifically developed for the analysis of Dutch dysarthric speech. The tool enables speech-language pathologists to obtain objective measures of these prosodic aspects in a standardized and fast way, and enables dysarthric speakers to practise their prosodic skills intensively without the presence of a speech-language pathologist being required.

Keywords: *dysarthria; speech rate; intonation; sentence modality; automated assessment; automated treatment*

I. INTRODUCTION

Dysarthria is a neurologic speech disorder characterized by abnormalities in the strength, speed, range, steadiness, tone, or accuracy of the movements of the speech musculature [1]. It is widely acknowledged that impairment of the prosodic aspects of speech (speech rate, intonation, and stress) is a perceptual hallmark of dysarthria [2]. Prosodic impairment can significantly hamper speech intelligibility [3], which has in its turn a negative influence on a speaker's daily life communicative exchanges and thus on quality of life in general. Therefore, assessment and treatment of prosody in individuals with dysarthria is clinically relevant.

Current assessment of prosody in dysarthria is mainly auditory-perceptual, which has some drawbacks. There are limitations to the reliability and validity of assessment procedures such as rating prosodic dimensions [4]. Moreover, alternative assessment procedures such as prosodic transcriptions or manual speech rate calculations are time-consuming [5].

Current treatment of prosody relies on modeling and auditory-perceptual feedback. Therefore, treatment is usually limited to therapy sessions in the presence of a speech-language pathologist (SLP). Motor learning research findings, however, suggest that speech motor skill training should be intensive for speech motor skill learning to occur [1].

We report progress on a speech-technology based prosody assessment and therapy program attempting to overcome the aforementioned limitations of current prosody assessment and

treatment practices. The tool is specifically intended for Dutch speaking individuals with dysarthria and at present allows for automatic measurement of speech rate (number of syllables uttered per second) and automatic intonation assessment (classification of an utterance according to its sentence modality: statement or question). In this way, the developed tool (1) enables SLPs to obtain standardized objective measures of speech rate and sentence modality and (2) enables individuals with dysarthria to practise these prosodic aspects intensively and autonomously, besides the classical speech therapy sessions in the presence of an SLP.

Development of automated prosody measures is not entirely new. Several algorithms for speech rate measurement of normal speech have already been reported, e.g. [6-8]. One speech rate algorithm has been tested on dysarthric speech [9], but has not resulted in a clinical application. Some research concerning sentence modality recognition has been reported for purposes such as automatic speech recognition [10] and prosodic assessment of language in hearing impaired children [11-12], but no appropriate methods are reported for the automatic assessment of intonation in dysarthric speech, except for [13], a precursor of the algorithm included in the currently presented tool. To our knowledge, the current work is the first automated prosody assessment and treatment program based upon algorithms specifically tuned to the analysis of dysarthric speech.

II. METHODS

A. Development of Speech Algorithms

For the development of the speech rate algorithm, speech samples of 28 healthy and 42 dysarthric speakers were collected. Per speaker, 18 sentence samples and 2 text samples were recorded. An SLP annotated all samples in Praat by placing boundaries between speech and pauses, counting the number of uttered syllables, and manually computing speech rate (SR) and articulation rate (AR). These annotated samples were used to evaluate and optimize the performance of a new, vowel detection based algorithm for automatically measuring speech rate in healthy and dysarthric speakers.

For the development of the intonation algorithm, a separate set of sentence samples of 20 healthy and 36 dysarthric speakers was composed, consisting of prosodic minimal pairs (e.g. Karen plays tennis./ Karen plays tennis?). Per speaker, three such utterances were recorded for each sentence modality

type. This enabled us to present the sentences in a randomized way to three SLPs, who classified each utterance perceptually as a statement or a question. Majority scores (2/3 or 3/3 agreement) served as a reference for training and testing of a newly developed algorithm for intonation classification in healthy and dysarthric speakers. To this end, the algorithm tracks the pitch movements in the last syllables of an utterance and also compares this end pitch to the speaker's pitch range.

B. Development of Software Application

Prior to the development of the graphical user interface for the assessment and therapy software tool, a center was consulted that specializes in advice on computers and communication for disabled people. In order to make the application useful for a wide range of dysarthric speakers (sometimes having visual and/or cognitive disorders, often elderly), it was decided to create a sober user interface. Special attention was given to adequate numerical feedback on prosodic performance after assessment and simple, effective, and motivating visual feedback on prosodic performance during therapy (see Fig. 1).

Therapy program content was developed along the lines of the assessment, providing training on sentence-level (for speech rate and intonation) and text-level (for speech rate) by means of various speech tasks (reading, imitation, choral speech, semi-spontaneous speech).

Clinical trials were organised to gather user feedback. The resulting input of five SLPs and four dysarthric speakers was used to further adjust the tool's interface to end-user needs. On two official occasions, an alpha version was presented to more than fifty SLPs, who provided further feedback on enhancing functionality and improving the interface.

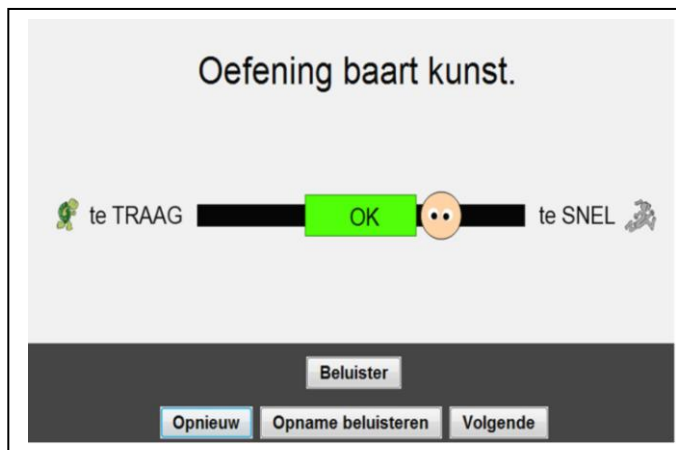


Figure 1. Visual feedback during speech rate exercise. The speaker read the sentence ('Practice makes perfect') too fast compared to the target speech rate.

III. RESULTS AND DISCUSSION

Algorithm performance was evaluated to see how well the developed objective measures correlate with perceptual measurements or judgments in normal speech and dysarthric speech.

TABLE I. AUTOMATED VERSUS PERCEPTUAL MEASUREMENT OF SPEECH RATE: PEARSON'S CORRELATION COEFFICIENTS

Speech Rate Algorithm	Correlation (r)	
	Normal speech	Dysarthric speech
text samples	0.96	0.96
sentence samples	0.83	0.87

For speech rate determination, high correlations were obtained for text-level samples, while correlations for sentence-level samples were somewhat lower (see table 1). This can be explained by the fact that the longer the time window for calculating speech rate, the less a measurement error will negatively influence the measure.

TABLE II. CLASSIFICATION ACCURACY OF INTONATION ALGORITHM

Intonation Algorithm	Accuracy of automated classification (%)	
	Normal speech	Dysarthric speech
statements	100.0	97.9
questions	96.7	90.3

For intonation classification, high classification accuracy was obtained for statements, while accuracy for questions was slightly lower (see table 2). Preliminary results seem to indicate that the algorithm generally performs equally well when classifying trained and new (non-trained) sentences. However, feedback on non-trained sentences is not always accurate. This is presumably due to the fact that these sentences have other phonetic and phonotactic characteristics than the trained sentences.

It needs to be mentioned that voiced speech is necessary in order not to compromise algorithm accuracy. This makes the tool unsuitable for dysarthric speakers suffering from aphonia. Generally, we can state that the tool is suitable in case of mild or moderate dysarthria, and less suitable in case of severe dysarthria.

The research reported here is a work in progress. Currently, the tool's therapeutic effectiveness is being studied in a group of twelve dysarthric individuals with Parkinson's disease, by means of a one-group pretest-posttest design. Preliminary results seem to indicate that a significant progress in speech intelligibility was obtained after three weeks of intensive therapy (five one-hour sessions per week) using the tool.

Release of a version inclusive of all desired functionality is scheduled for spring 2013. This will enable us to conduct a usability study in the near future. Further research will focus on optimization of both algorithms and creation of an algorithm for the detection of sentence stress in dysarthric speech. The current state of this research project can also be followed on <http://catris.etro.vub.ac.be/index.html>.

ACKNOWLEDGMENT

The authors were able to conduct this research by means of research grant IWT-TBM 080662 from the Flemish Agency for Innovation by Science and Technology (Vlaams Agentschap

voor Innovatie door Wetenschap en Technologie - IWT) and by means of a scholarship from the Flemish Interuniversity Council (Vlaamse Interuniversitaire Raad - VLIR). None of the authors has any relevant financial relationships with any commercial interests.

REFERENCES

- [1] J.R. Duffy, *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 2nd ed. Saint Louis, MO: Elsevier Mosby, 2005.
- [2] R. Patel, "Assessment of Prosody," in *Assessment of Motor Speech Disorders*, A. Lowit and R.D. Kent, Eds. San Diego, CA: Plural Publishing, 2011, pp. 75-95.
- [3] M. De Bodt, M.E. Hernandez-Diaz Huici, and P.H. Van de Heyning, "Intelligibility as a linear combination of dimensions of dysarthric speech," *Journal of Communication Disorders*, vol. 35(3), pp. 283-292, 2002.
- [4] R.D. Kent, "Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders," *American Journal of Speech-Language Pathology*, vol. 5(3), pp. 7-23, 1996.
- [5] G. Van Nuffelen, "Speech Prosody in Dysarthria: Assessment and Remediation," in *Speech Prosody in Atypical Populations*, V. Stojanovik and J. Setter, Eds. Guildford: J&R Press, 2011, pp. 147-167.
- [6] N. Morgan, and E. Fossler-Lussier, "Combining multiple estimators of speaking rate," *IEEE ICASSP*, Seattle, WA, vol. 2, pp. 729-732, May 1998.
- [7] D. Wang, and S.S. Narayanan, "Robust speech rate estimation for spontaneous speech," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15(8), pp. 2190-2201, November 2007.
- [8] N.H. De Jong, and T. Wempe, "Praat script to detect syllable nuclei and measure speech rate automatically," *Behavior Research Methods*, vol. 41(2), pp. 385-390, 2009.
- [9] M. Mujumdar, *Estimation of the Number of Syllables using Hidden Markov Models and Design of a Dysarthria Classifier using Global Statistics of Speech*. PhD thesis. Laramie, WY: University of Wyoming, 2006.
- [10] H. Wright, "Automatic utterance type detection using suprasegmental features," *Proceedings of the International Conference on Spoken Language Processing*, paper 0575, 1998.
- [11] F. Ringeval, J. Demouy, G. Szaszák, M. Chetouani, L. Robel, J. Xavier, D. Cohen, and M. Plaza, "Automatic intonation recognition for the prosodic assessment of language impaired children," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19(5), pp.1328-1342, July 2011.
- [12] P. Král, J. Klečková, and C. Cerisara, "Sentence modality recognition in French based on prosody," *World Academy of Science, Engineering and Technology*, vol. 8, pp. 677-680, 2007.
- [13] D. Torres, T. Dekens, H. Martens, G. Van Nuffelen, M. De Bodt, W. Verhelst, and C.A. Ferrer, "Sentence modality recognition in dysarthric speech," *Proceedings of MAVEBA 2011*, Florence, Italy, pp. 71-74, August 25-27, 2011.