

DIA: A TOOL FOR OBJECTIVE INTELLIGIBILITY ASSESSMENT OF PATHOLOGICAL SPEECH

C. Middag¹, J.P. Martens¹, G. Van Nuffelen² and M. De Bodt²

¹Department of Electronics and Information Systems, Ghent University, 9000 Ghent, Belgium

²Antwerp University Hospital, University of Antwerp, 2650 Edegem, Belgium

Abstract: Intelligibility is generally accepted to be a very relevant measure in the assessment of pathological speech. In clinical practice, intelligibility is measured using one of the many existing perceptual tests. These tests usually have the drawback that they employ unnatural speech material (e.g. nonsense words) and that they cannot fully exclude errors due to the listener's bias. This raises the need for an objective and automated tool to measure intelligibility. Here, we present the Dutch Intelligibility Assessment (DIA), an objective tool that aids the speech therapist in evaluating the intelligibility of persons with pathological speech. This tool will soon be made publicly available.

Keywords : objective intelligibility assessment, pathological speech, speech therapy.

I. INTRODUCTION

Communication is getting increasingly important in our society. People with communication disorders often suffer other social discomfort as well. Therefore, follow-up of these patients in order to improve their pronunciation is becoming increasingly important. An important measure in the assessment of communication efficiency is intelligibility. This is often determined in a perceptual way, which is subjective in nature. Consequently, there is a growing need for an automated, and thus objective method for measuring intelligibility.

Previous software packages have been developed to measure intelligibility of English patients suffering from dysarthria [1] and [2] describes a system doing the same for German laryngectomees and children with cleft lip and/or palate. The former is based on a measure called goodness of fit of the alignment between the uttered speech and the target speech, while the latter uses word accuracy rate, after doing speech recognition on the uttered speech.

We present the Dutch Intelligibility Assessment (DIA), a tool to assist speech therapists when dealing with patients suffering from pathological speech, which is based on a novel methodology, described in [3]. The method underlying the DIA tool extracts phonemic, phonological and context-dependent phonological features from automatic speech alignment on the basis of acoustic models that were trained on normal speech.

Based on those features, intelligibility is predicted using a compact model that can be trained on pathological speech samples. The experimental evaluation of the system shows standard errors between perceived and computed intelligibilities lower than 8%. This is a sufficiently strong basis for the development of an automated version of the Dutch Intelligibility Assessment.

II. THE PERCEPTUAL DIA TEST

The –initially subjective– test we have automated is the Dutch Intelligibility Assessment (DIA) test [4]. This test consists of 50 consonant-vowel-consonant (CVC) words, mostly nonsense but well pronounceable words. These 50 words are divided into three subtests: one testing the Dutch consonants in the initial position, one in the final position and the last one testing the vowels and diphthongs in the middle position in the word. To avoid guessing by the listener, there are 25 variants of each subtest, of which one is chosen at random for each execution of the test. The perceptual intelligibility score is then calculated as the percentage of tested phonemes which are correctly identified. This test is proven to be highly reliable (an interrater correlation of 0.91 and an intrarater correlation of 0.93 [4,5]).

III. THE COMPUTERIZED DIA TOOL

Within the framework of the SPACE¹ project, this perceptual DIA test has been automated, as described in [6]. While the perceptual test only uses the 50 tested phonemes, the computerized version takes every phoneme of the 50 words into account. All uttered speech is lined up against the target words using forced alignment of two automatic speech recognizers (ASRs). This results in three feature sets: phonemic features, phonological features and context-dependent phonological features. The phonemic features describe how well on average the Dutch phonemes are recognized by the used ASRs, while the phonological features describe how well a phonological feature can be realized by the speaker. The context-dependent phonological features point to transitions between two articulatory positions. These feature sets are then used in a simple regression model to predict the intelligibility of the speaker.

¹ <http://www.esat.kuleuven.be/psi/spraak/projects/SPACE/>

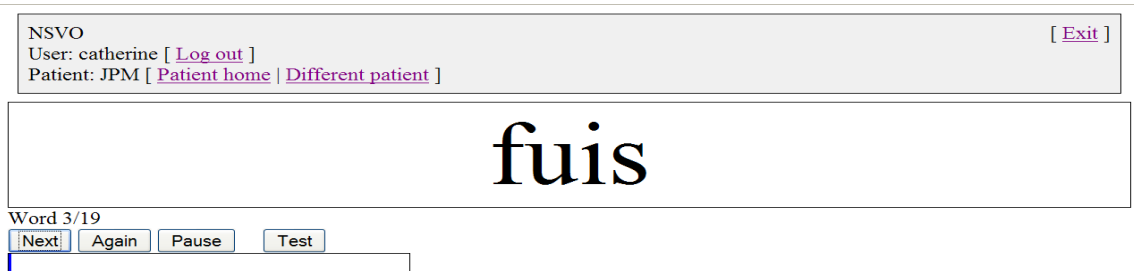


Figure 1 Screenshot of the recording environment of the tool. The patient is presented with a sequence of words, which are displayed one after the other. These words are automatically recorded by the tool for later analysis.

Different models have been designed: a general model, as well as pathology-specific models for people with hearing impairment, dysarthria and laryngectomy. We recently also added a model for children with cleft lip and palate. As shown in [3,6], the correlations between the computed intelligibility scores and the perceptual scores are about as high as the interrater reliability, which means the automated version can compete with the human judging. Moreover, the DIA tool could be a more objective and less time-consuming way for the speech therapist to administer the test.

IV. TOOL DESCRIPTION

Our purpose was to design a user-friendly, easily available tool which does not require a complex setup to administer the test. To use the DIA tool, the user only needs a PC or laptop with a web browser, a head set and sound card, and an up-to-date Java runtime environment. The tool works in a client/server environment and can be used both in online or offline mode.

Once a user has an account, patients can be added and edited. As we respect the privacy of the patients, every user can only view recordings of its own patients. When a patient is added, the user can start the test. We advice to do a microphone test first, to be assured that the recording quality is well enough and the microphone is in the right position (e.g. not too close to the mouth). When starting the test, a sequence of words is presented to the patient (Fig. 1). Each of these is recorded as a separate .wav file, which is stored for subsequent analysis.

When the recording is finished, the speech therapist can analyze the recordings by listening to every word and filling in the missing phoneme (Fig. 2). This results in a perceptual score and a report displaying the nature of the errors, e.g. wrong place/manner of articulation, as described in [4,5]. Every recording can be judged by several listeners, which can easily be added in the recording menu.

In a final step, the user can also run an automatic analysis. This step results in an objective intelligibility

score, as well as a number of statistics of the analysis (Fig. 3). These statistics display the speech profile of the current patient, compared to normal speakers, as well as a number of well-defined pathologies.

V. TOOL VALIDATION

To validate the tool, a master student recorded 33 laryngectomees, 19 hearing impaired, and 9 dysarthric patients. The recording settings were not always ideal and sometimes a lot of background noise could be noticed. Every patient performed the test, which was recorded using our DIA tool.

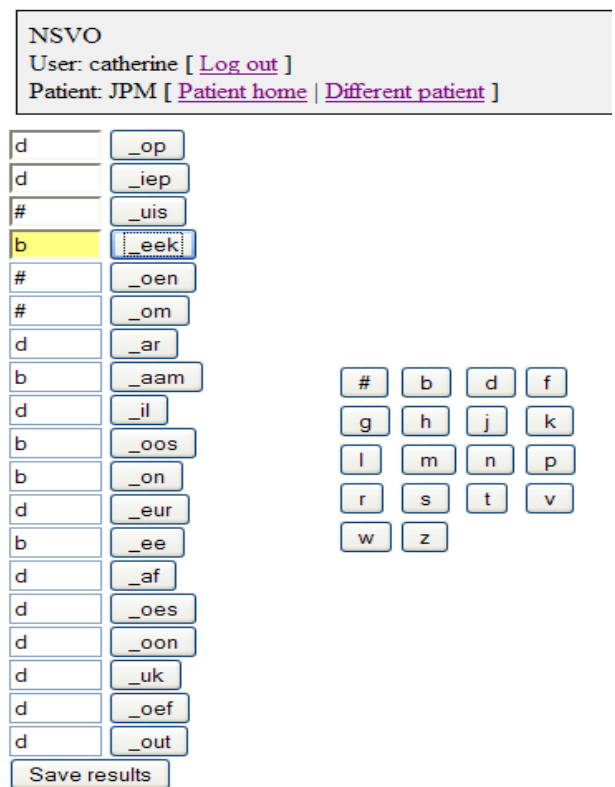


Figure 2 Perceptual analysis of the recordings. When clicking on the button, the corresponding .wav file is played, and the listener can fill in the missing part.

Automatic score

94%

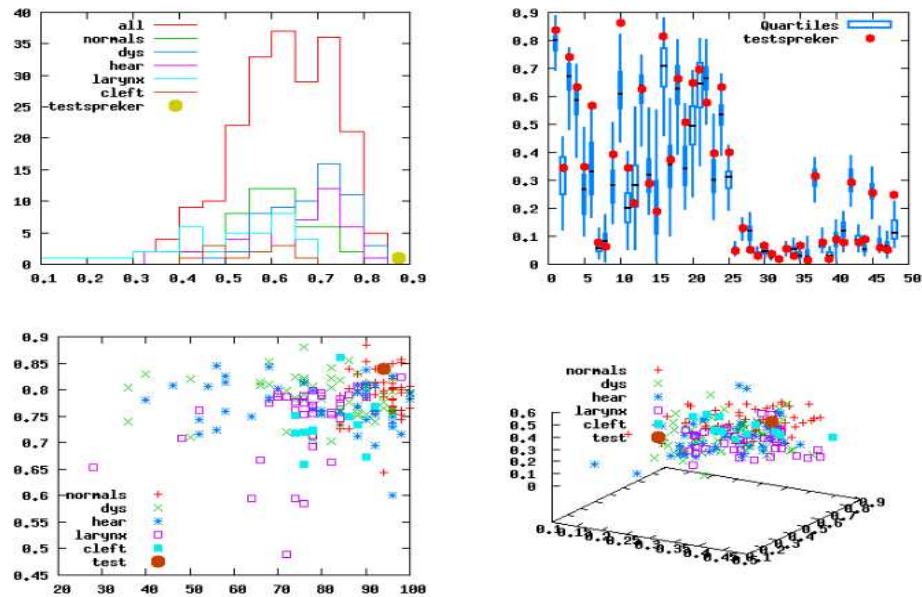


Figure 3 A part of the final report, showing the results of the automatic analysis. Here, an overall score for the objective intelligibility is calculated (upper left), as well as a number of statistics, showing a.o. how the current patient performs with respect to a normal speaker (label “testpreker”) as well as a number of well-defined speech pathologies.

Apart from the objective score calculation, the subjective evaluation of the speech intelligibility was performed by two professional listeners.

The interrater agreement between the two listeners was measured using the Pearson correlation coefficient between their scores and reached values as high as 94%. The Pearson correlation between the mean of the listener’s scores and the objective scores reached 90%, which is almost as good as the interrater agreement.

VI. FUTURE WORK

These results are very promising and reveal that an objective evaluation of pathological speech can indeed be useful in the clinical practice. In a next step, we will investigate the possibility of replacing the nonsense words by more natural speech such as existing words or even phrases. We are also working towards a more profound articulatory assessment, which can then lead to the determination of an appropriate therapy for every patient.

REFERENCES

[1] J. N. Carmichael, “Introducing Objective Acoustic Metrics for the Frenchay Dysarthria Assessment Procedure”, PhD Thesis, 2007, University of Sheffield.

[2] A. Maier, T. Haderlein, U. Eysholdt, F. Rosanowski, A. Batliner, M. Schuster and E. Nöth, “PEAKS – A system for the automatic evaluation of voice and speech disorders”, in *Speech Communication* vol. 51, 2009, pp. 425-437

[3] C. Middag, J.P. Martens, G. Van Nuffelen, and M. De Bodt, “Automated Intelligibility Assessment of Pathological Speech Using Phonological Features,” in *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 629030, 9 pages, 2009. doi:10.1155/2009/629030

[4] M. De Bodt, C. Guns, and G. V. Nuffelen, NSVO: *Nederlandstalig SpraakVerstaanbaarheidsOnderzoek*, Vlaamse Vereniging voor Logopedisten, Herentals, Belgium, 2006.

[5] G. Van Nuffelen, M. De Bodt, C. Guns, F. Wuyts, and P. Van de Heyning, “Reliability and clinical relevance of segmental analysis based on intelligibility assessment,” *Folia Phoniatria et Logopaedica*, vol. 60, no. 5, pp. 264–268, 2008.

[6] C. Middag, G. Van Nuffelen, J. P. Martens, and M. De Bodt, “Objective intelligibility assessment of pathological speakers,” in *Proceedings of the International Conference on Spoken Language Processing (Interspeech '08)*, pp. 1745–1748, Brisbane, Australia, September 2008.