

RESEARCH REPORT

Automatic boost articulation therapy in adults with dysarthria: Acceptability, usability and user interaction

Viviana Mendoza Ramos^{1,3}  | Juan C. Vasquez-Correa²  | Rani Cremers⁴ |
Leen Van Den Steen¹ | Elmar Nöth² | Marc De Bodt^{1,3,4} | Gwen Van Nuffelen^{1,3,4}

¹ Department of Otorhinolaryngology, Head and Neck Surgery and Communication Disorders, University Hospital of Antwerp, Edegem, Belgium

² Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

³ Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Antwerp, Belgium

⁴ Faculty of Medicine and Social Health Sciences, University of Ghent, Ghent, Belgium

Correspondence

Viviana Mendoza Ramos, Department of Otorhinolaryngology, Head and Neck Surgery and Communication Disorders, University Hospital of Antwerp, Edegem, Belgium.

Email: Viviana.MendozaRamos@uza.be

ABSTRACT

Background: Imprecise articulation has a negative impact on speech intelligibility. Therefore, treatment of articulation is clinically relevant in patients with dysarthria. In order to be effective and according to the principles of motor learning, articulation therapy needs to be intensive, well organized, with adequate feedback and requires frequent practice.

Aims: The aims of this pilot study are (1) to evaluate the feasibility of a virtual articulation therapy (VAT) to guide patients with dysarthria through a boost articulation therapy (BART) program; (2) to evaluate the acoustic models' performance used for automatic phonological error detection; and (3) to validate the system by end-users from their perspective.

Methods & Procedures: The VAT provides an extensive and well-structured package of exercises with visual and auditory modelling and adequate feedback on the utterances. The tool incorporates automated methods to detect phonological errors, which are specifically designed to analyse Dutch speech production. A total of 14 subjects with dysarthria evaluated the acceptability, usability and user interaction with the VAT based on two completed therapy sessions using a self-designed questionnaire.

Outcomes & Results: In general, participants were positive about the new computer-based therapy approach. The algorithm performance for phonological error detection shows it to be accurate, which contributes to adequate feedback of utterance production. The results of the study indicate that the VAT has a user-friendly interface that can be used independently by patients with dysarthria who have sufficient cognitive, linguistic, motoric and sensory skills to benefit from speech therapy. Recommendations were given by the end-users to further optimize the program and to ensure user engagement.

Conclusions & Implications: The initial implementation of an automatic BART shows it to be feasible and well accepted by end-users. The tool is an appropriate solution to increase the frequency and intensity of articulation training that supports traditional methods.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *International Journal of Language & Communication Disorders* published by John Wiley & Sons Ltd on behalf of Royal College of Speech and Language Therapists

What this paper adds

What is already known on the subject

- Behavioural interventions to improve articulation in patients with dysarthria demand intensive treatments, repetitive practice and feedback. However, the current treatments are mainly limited in time to the interactive sessions in the presence of speech–language pathology. Automatic systems addressing the needs of individuals with dysarthria are scarce. This study evaluates the feasibility of a VAT program and investigates its acceptability, usability and user interaction.

What this paper adds to existing knowledge

- The computer-based speech therapy approach developed and applied in this study intends to support intensive articulation training of patients with dysarthria. The virtual speech therapy offers the possibility of an individualized and customized therapy programme, with an extensive database of exercises, visual and auditory models of the target utterances, and providing adequate feedback based on automatic acoustic analysis of speech.

What are the potential or actual clinical implications of this work?

- The automatic BART overcomes the limitation in time of face-to-face traditional speech therapy. It offers patients the opportunity to have access to speech therapy more intensively and frequently in their home environment.

INTRODUCTION

Dysarthria is a neurogenic motor speech disorder caused by impaired muscular control due to damage to the central or peripheral nervous system (Darley *et al.* 1969, 1975). The impairment may affect the mechanisms involved in speech production and consequently cause deficits in speech intelligibility. Imprecise speech sound articulation is a hallmark of dysarthria (Tjaden 2007, Weismer 2007). Research in this domain has shown that articulation is the most essential dimension of speech contributing to the overall speech intelligibility in dysarthria (Kent 1992, De Bodt *et al.* 2002). Deficient articulatory movements are linked with reduced acoustic contrast for consonants and vowels, which leads to reduced intelligibility (Tjaden 2007, Tykalova *et al.* 2017). Even when neurologic bases differ among different types of dysarthrias, the impact on articulatory movements and acoustic outcomes is in general comparable (Tjaden 2007). Reduced speech intelligibility has a major psychosocial impact. Daily life communication gets less evident, satisfactory or even impossible for subjects with dysarthria, affecting general quality of life (Dykstra *et al.* 2007, Dickson *et al.* 2008). Therefore, behavioural interventions to improve articulation and, consequently,

speech intelligibility are necessary for individuals with dysarthria.

There is scientific and clinical evidence that improving articulatory skills demands intensive and frequent therapy (Duffy 2019, Rosenbek and Jones 2009). There is growing evidence that the content of speech therapy must meet the principles of motor learning (Duffy 2019). Key elements of motor learning in clinical neurorehabilitation are intensive treatment, repetitive practice and sensory feedback (Kleim *et al.* 2003, Garvey *et al.* 2007, Kleim and Jones 2008, Maas *et al.* 2008). Extensive research in the speech therapy domain suggests that providing large amounts of practice over a shorter period can lead to better results/outcomes for adults with communication disorders (Bhogal *et al.* 2003, Fox *et al.* 2006). A more recent study demonstrated that boost articulation therapy (BART), in which participants perform intensive articulatory drill and minimal pairs exercises for 5 consecutive days, has a significant positive effect on speech intelligibility in patients with dysarthria (Mendoza *et al.* 2021). However, the current treatments of articulation are mainly limited in time to the interactive sessions in the presence of speech–language pathology (SLP). A computer-based speech therapy approach would overcome that limitation



offering patients the opportunity to practise more intensively and frequently to reduce the number of fatiguing and time-consuming transfers to the rehabilitation centre and to have access to speech therapy in their home environment. Virtual speech–language therapy would also provide patients with the possibility of an individualized and customized therapy programme. However, this virtual therapy program only targets those persons with dysarthria who have sufficient cognitive, linguistic, motoric and sensory (auditory and visual) skills to benefit from speech therapy (Weismer 2007).

Virtual therapy technologies are not entirely new; however, only a few attempts have been made for adults with motor speech disorders (Lee 2019). Most of the reported studies using virtual therapy focus on adults with aphasia (e.g., Abad *et al.* 2013, Thompson *et al.* 2010, Van Vuuren and Cherney 2014, Cherney *et al.* 2008, Lee *et al.* 2009, Kalinyak-Fliszar *et al.* 2015, Marshall *et al.* 2016, Amaya *et al.* 2018) and children with hearing impairment (Chaisanit *et al.* 2010, Eriksson *et al.* 2005, Massaro and Light 2004, Silva *et al.* 2012). Automatic systems addressing the needs of individuals with dysarthria are scarce (Lee 2019). The designed and tested experimental models include the single case study of Cole *et al.* (2007), which investigated the effects of Lee Silverman Voice Treatment (LSVT) with a virtual therapy system; qualitative evaluation of the system by the user shows positive results. The study of Palmer *et al.* (2007) compared the effects of computerized and traditional therapy for speakers with longstanding dysarthria. The computer-based program consists of games that target the accurate production of individual speech sounds and exercises for practising sounds in words, phonation, pitch and volume, allowing independent training with feedback. This approach was demonstrated to be as effective as traditional therapy improving the speech of that specific group. The study of Saz *et al.* (2009), which evaluated a semi-automated tool that includes different games for children and young adults, considers phonation, phonetic articulation and language understanding for Spanish speakers. The feedback given about the system by the therapist and end-users was mainly positive. The study of Orozco-Aroyave *et al.* (2020) introduced a mobile application for monitoring and evaluation of patients with Parkinson's disease. The app considers different motor aspects such as speech production, limb movements and finger tapping. It allows patients to track their progression and can be used as a personal health assistant. To our knowledge, there is no publicly available virtual articulation tool providing BArT with adequate feedback, especially not for Dutch-speaking subjects with articulatory deficits.

The present pilot study aims to evaluate the feasibility of a VAT for individuals with dysarthria. The goal of the

system is to guide the patient through an intensive treatment programme based on the principles of motor learning and relying on speech technology for adequate feedback. It is intended to improve articulation and consequently pronunciation, intelligibility and communication abilities.

The study also focuses on the evaluation of the system by end-users. The purpose is to investigate the acceptability, usability and user interaction with computer-assisted articulation therapy. The stakeholder's opinion is particularly valuable to ensure that the new virtual intervention provides a suitable and engaging treatment. Some important aspects to consider are the participants' ability to interact independently with the virtual articulation trainer; whether they find the user interface attractive, supportive and clear, and if they find the material to be useful. It is also important to know how end-users experience the visual feedback and their general perception of the program.

Therefore, the main goals of this study are the design and development of a new computer-based therapy program that supports intensive articulation training of patients with dysarthria, the evaluation of acoustic models' performance used for automatic phonological error detection, and a general validation by stakeholders of the usability, acceptability and interaction with the automatic system.

MATERIALS AND METHODS

Participants

The developed program targets patients with dysarthria who have sufficient cognitive, motor and sensory (auditory and visual) skills to work with a computer and/or those who can rely on assistance.

In order to participate in the study, the patients had to be adults (> 18 years of age) with a diagnosis of dysarthria due to neurogenic origin (confirmed by an experienced SLP and neurologist), native speakers of Dutch and with a score < 90% on the Dutch phoneme Intelligibility Assessment (DIA) (De Bodt *et al.* 2006).

Participants were excluded from the study if they had serious visual or auditory problems, insufficient computer skills to use the software program independently and if an intensive articulation therapy was followed less than 2 months before the study in order to exclude possible effects of the training in the baseline of the intelligibility assessment.

Patients were recruited by an experienced SLP at the Noorderhart Rehabilitation & Multiple Sclerosis Center Pelt in Hasselt by convenience sampling. A total of 14 subjects with dysarthria participated in this pilot study; their severity level was determined perceptually by the SLP.

TABLE 1 Characteristics of the participants

Participant	Sex	Age (years)	Neurological pathology	Type of dysarthria	Severity dysarthria	DIA score (%)	SHI score
1	M	59	MS	Flaccid	Moderate	80%	60
2	M	76	Stroke	Flaccid	Moderate	82%	18
3	F	76	Parkinson's disease	Hypokinetic	Mild	90%	32
4	M	58	MS	Mixed	Mild	68%	20
5	M	70	Stroke	Flaccid	Moderate	60%	39
6	M	67	Stroke	Flaccid	Mild	90%	10
7	M	47	Friedreich ataxia	Ataxic	Mild	82%	17
8	M	57	MS	Flaccid	Moderate	62%	41
9	M	29	TBI	Flaccid	Mild	84%	13
10	F	76	Parkinson's disease	Hypokinetic	Mild	88%	38
11	F	67	Stroke	Flaccid	Mild	76%	25
12	F	92	Stroke	Ataxic	Mild	88%	51
13	F	66	MS	Spastic	Mild	90%	23
14	M	67	Stroke	Spastic	Moderate	76%	1

Notes: M, male; F, female; MS, multiple sclerosis; TBI, traumatic brain injury; DIA, Dutch phoneme Intelligibility Assessment. Score of the Speech Handicap Index (SHI) (score range: 0–60; score < 14 no impact, 14–22 light, 23–31 moderate, > 31 severe impact).

Table 1 displays the subjects' characteristics. The Speech Handicap Index (SHI) (Van den Steen *et al.* 2011) was administered to each participant; this instrument is a self-evaluation of quality of life and quantifies the biopsychosocial impact of the speech disorder.

Ethics permission

Ethical approval for this study was obtained by the Regional Committee of Medical and Health Research Ethics (B300201837099). All participants agreed voluntarily to participate in the study and signed an informed consent form.

Software

Design criteria

The VAT was designed for patients with articulation deficits in which a virtual therapist, acting like a real SLP, guides the patient through an intensive treatment program to improve articulation and consequently speech intelligibility. The developed software runs as a desktop application and can be installed on any computer. The VAT has multiple roles and functions: motivating and stimulating the patient, giving clear instructions, analysing the target

utterance, and providing adequate feedback based on automatic acoustic analysis of speech.

The development of the VAT system consisted of three main steps: (1) the development of an extensive database of exercises for articulation training; (2) the implementation of an algorithm for phonological error detection in order to provide the necessary feedback to the patients; and (3) the embedding of the set of new developments in an attractive, patient-friendly user interface.

Extensive database of exercises and intervention design

The integrity of speech sounds is a major factor in determining intelligibility (Kent 1992), thus the direct treatment of articulation at the segmental level of speech is the aim of this therapy program. This approach intends to reduce the speech impairment, and its effects may be also generalized to connected speech. A strong correlation between word intelligibility and sentence intelligibility was also found in previous studies (Yorkston and Beukelman 1978, Weismer *et al.* 2001).

The content of speech material consists of two different: intensive training programs to improve articulation and phoneme intelligibility in persons with dysarthria. Based on the principles of motor learning, the training programs comprise articulatory drill and minimal pairs.

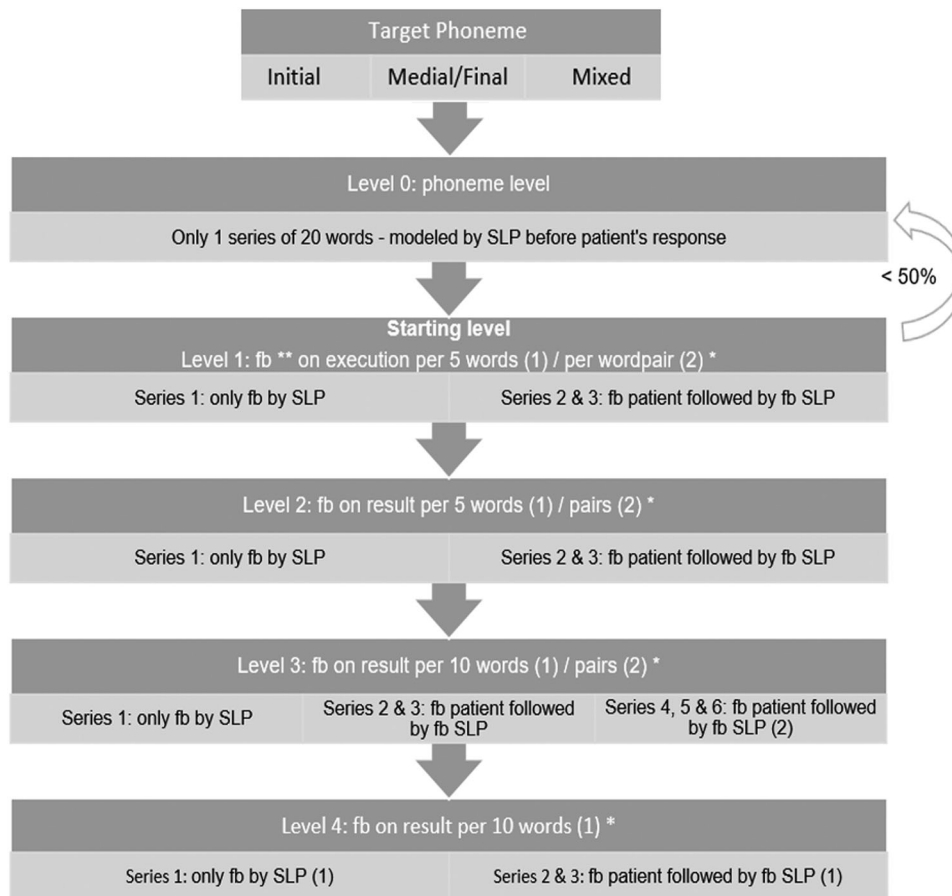


FIGURE 1 The intervention design (Mendoza *et al.* 2021). The components labelled (1) are exclusively intended for the intervention articulatory drill. The components labelled (2) are intended for the intervention of minimal pairs. When nothing is included, the parts are common for both interventions. *SLP scores each word by means of a scoring form; and **feedback.

There are four levels of difficulty in each program, structured per target position of the consonant, in initial, medial and final position in the words. This allows the patient to go through intensive therapy. In a previous study, this particular approach (figure 1) resulted in significant improvements of intelligibility at different levels of speech in a group of adults with dysarthria (Mendoza *et al.* 2021). The extensive, well-structured and hierarchical practice package of exercises recommended in that previous study is implemented in the VAT.

The exercises only include a series of mono- and bisyllabic words. These types of words are usually not problematic to understand or to read for patients with mild cognitive impairments. Most of the selected words are meaningful and emphasize the correct pronunciation of the target phoneme, making the therapy more effective and goal oriented. The stimuli were carefully selected by experienced SLPs from a traditional articulation programme 'Articulation in Practice' (Huybrechts *et al.* 1999) and from the official word inventory of Dutch (Renkema 1995) to increase the number of stimuli.

Algorithm selection and modifications

The algorithm used in the program for phonological error detection derived from the acoustic signal is based on existing models (Vásquez-Correa *et al.* 2019) using state-of-the-art gated recurrent neural networks (GRNN). It has been shown to be accurate at evaluating the dysarthria severity of patients with motor speech disorders, such as patients affected with the neurodegenerative Parkinson's disease (Orozco-Arroyave *et al.* 2020, Miller *et al.* 2020).

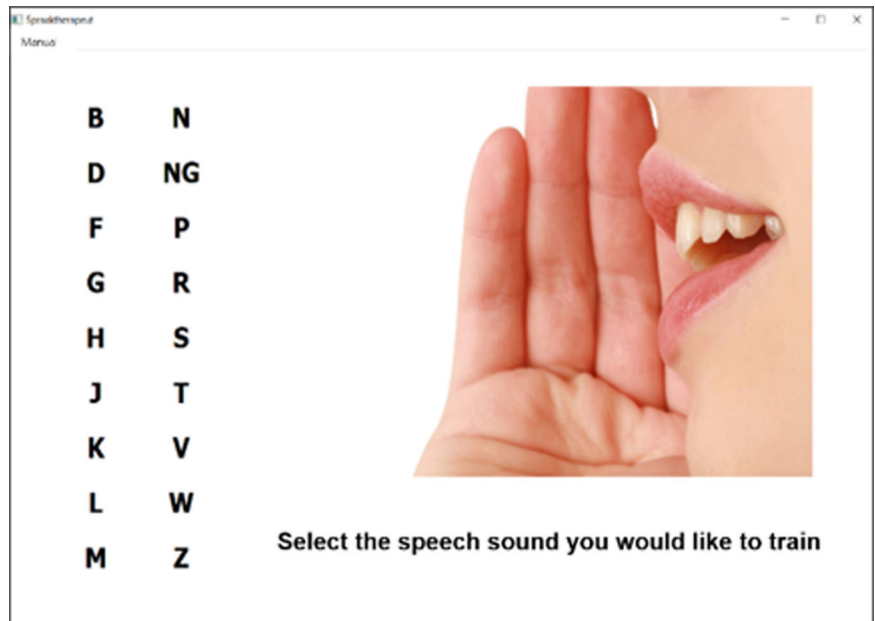
The algorithm is a statistical approach to the human perception process and is based on the speech-production mechanisms. It extracts the log Mel-filterbank energy features from the acoustic signal and maps them into phonological features. Individual speech frames are represented by a vector with information about place and manner of articulation. This vector captures the posterior probability of a speech frame to belong to one or more phonological classes (e.g., plosive, nasal, bilabial). The model was originally designed for

TABLE 2 Consonants place and manner of articulation (Verhoeven 2005)

	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Glottal
Plosive	p b		t d		(c)	k (g)		(?)
Nasal	m	(ɱ)	n		(ɲ)	ŋ		
Trill			(r)				R	
Fricative		f v	s z	(ʃ) (ʒ)		x ɣ		ɦ
Approximant	w				j			
Lateral approximant			l		(ʎ)			

Note: Voiced consonants are on the right side.

FIGURE 2 Sound selection. [Colour figure can be viewed at wileyonlinelibrary.com]



Spanish, and most of the extracted characteristics were related to vowels. Therefore, modifications were necessary in order to reliably characterize consonants, and extended to detect the Dutch-specific phonological characteristics.

The combination of place and manner of articulation and voiced characteristics were used to describe the consonants in Dutch. Hence, the algorithm was retrained using the language-specific phonological features (Verhoeven 2005) (table 2) along with the read speech part of the Corpus Spoken Dutch (Oostdijk and Broeder 2003), which is an extensive corpus (903,043 words) of contemporary standard Dutch as it is spoken nowadays in the Netherlands and Dutch-speaking part of Belgium. With the refinement and extension of the models, trained on healthy speech and capable to recognize all the phonological features, it can be used to detect more specific articulation errors/deficits in subjects with dysarthria. Accuracy, precision (positive predictive value), recall (sensitivity) and *F*-score (harmonic mean of the precision and recall) performance measures to evaluate the models are offered in the results section.

The newly trained algorithm extracts the phonological posterior probabilities from the acoustic signal and it is used to identify phonological errors present during the articulation training. However, the system does not make any decision with respect to the type and degree of error at the phoneme level (if the error was an addition, substitution, omission or distortion).

User interface

The training program (VAT) has three stages. In the first, the user selects the target sound that he or she will train (figure 2). Once the selection is made, a video instruction with the description and examples of the correct production of the sound (visual model) is given. This support is provided with frontal and lateral video recordings that show the correct way of sound production in an easy, friendly and interpretable way (figure 3). Finally, auditory modelling of the words is also offered when the set of exercises starts. The patient repeats the target words, and the system will automatically capture and analyse the signal.

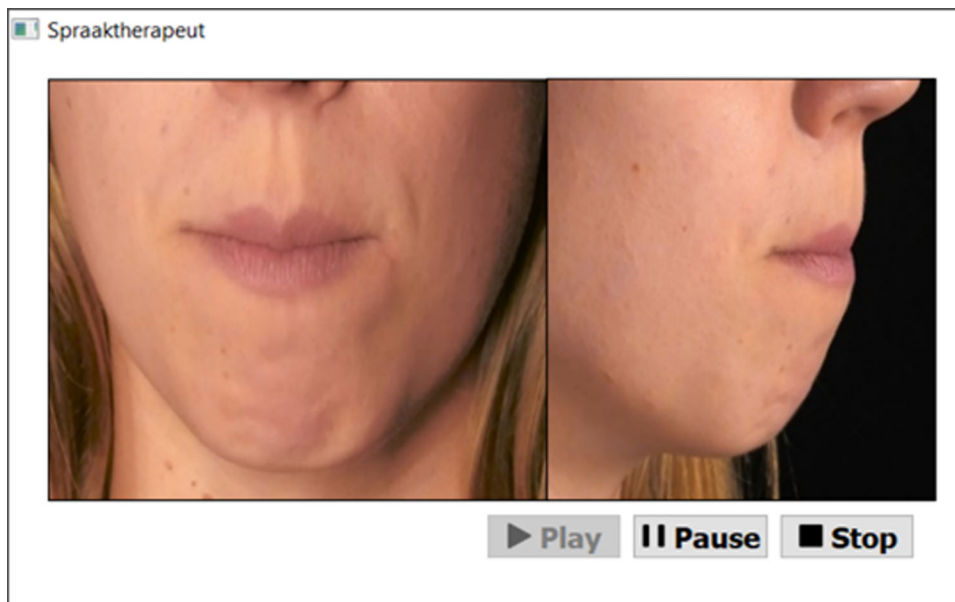


FIGURE 3 Visual modelling. [Colour figure can be viewed at wileyonlinelibrary.com]

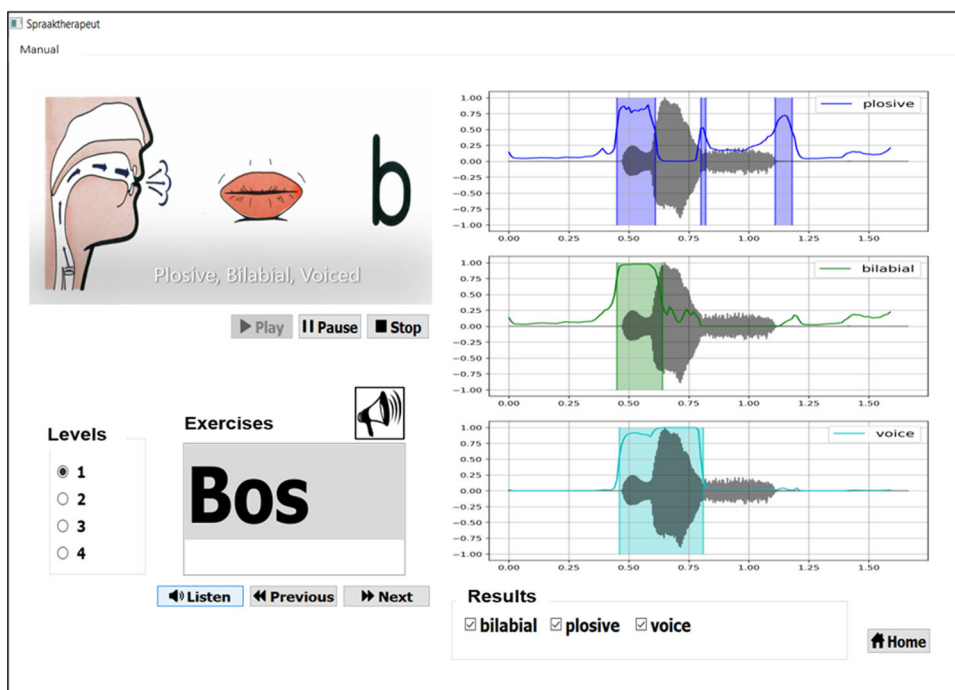


FIGURE 4 Auditory modelling, exercises and feedback. [Colour figure can be viewed at wileyonlinelibrary.com]

Subsequently, detailed visual feedback on the articulatory precision is provided. The graphs show whether or not the distinctive features of the target sound were present (figure 4). This detailed information, which is generally difficult to impossible to detect and quantify by the human ear, allows patients to work more precisely in the improvement of the articulatory movements. Misclassification or

false detections of distinctive features could occur in the case of present background noise; therefore, it is recommended that patients practise in a quiet room and are encouraged to use a head-mounted condenser microphone.

For the general design of the user interface, the font, font size and button sizes are big, simple and clean, intended for elderly and patients with brain damage.

Instruments to measure user interaction, acceptability and usability

After two therapy sessions of 45 min each, all the participants evaluated the computer-based therapy program. Usability, acceptability and interaction with the software were assessed through a questionnaire.

A questionnaire designed specifically for this study was completed by end-users after two sessions of practising with the VAT. The survey had the intention of evaluating software usability and acceptability from the user's point of view. The questionnaire was designed based on other questionnaires that addressed similar software-related topics, such as the Software Usability Measurement Inventory (SUMI) (Kirakowski and Corbett 1993), the General Computer Skill Questionnaire (Klinkenberg *et al.* 2004) and the ASISTO questionnaire (Deruyter and Staessens 2018); the items were restructured focusing on the VAT.

The set of questions addresses important topics that should be considered in order to succeed in future user engagement. It has a total of 32 items and is divided into five sections:

- Usability and operability: this section is designed to evaluate the patient's perception of the software functionality as a tool, if they feel comfortable working with it, if they can interact independently with it and if they find it friendly and easy to use, in general.
- Program and layout: this section focuses on how the users experience the visual and auditory support received during them practising and how attractive and clear the layout is perceived.
- Exercises: this section is designed to evaluate the user's perception of the usefulness of the training therapy and if they feel that the selected set of exercises is helpful for improving articulatory skills and consequently speech intelligibility.
- Feedback: this section is designed to evaluate the perception of the visual feedback that is given during therapy.
- Experience in the use of multimedia: with this section we would like to have the tendency of multimedia usage in daily life. It will give some insights into how likely or often we can expect the use of the software.

Each section contains several statements that participants have to score using a five-point Likert scale, where a score of 1 corresponds to 'completely disagree' and a score of 5 corresponds to 'completely agree'. The number of statements varies among sections; therefore, an average score is necessary for each section. Thereby, the questionnaire can have a maximum score of 25. The survey has also a section where the user might add suggestions or comments in order to improve the engagement and usability of the soft-

ware. The survey is applied after two therapy sessions. For an overview of the questionnaire, see [appendix A](#).

Training sessions

Before the start of the training, the goal of the virtual therapy and the steps of the program were explained in detail to every subject who participated in the study.

Training session 1

At the beginning of the first session, the DIA and SHI were conducted, followed by the supervised training moment. The selection of the target speech sounds to be trained was based on the error analysis of the DIA. The speech therapist started the computer and software and explained to the patient how to select the target speech sound and the level. Thereafter, it was demonstrated how to obtain auditory modelling of a training item, how to click to the next or preceding item, how to listen again to the patient's own speech production and how to go back to start. It was explained in detail how to understand and manage the given feedback. After the first session, all subjects received a printed version of the manual to consult at home if necessary.

Training session 2

The second session took place after a maximum of 2 days following the first session. During this session, patients worked completely autonomously with the VAT. It was expected to train with the same target sounds at the same level as during the first session. They were supposed to understand and manage the visual feedback in order to adjust their articulation. The questionnaire was administered at the end of this session. The supervisor was present for assistance if necessary.

Statistical analysis

Statistical analysis of the data was performed with the software program IBM Statistical Package for the Social Sciences v. 25 (IBM SPSS 25). The distribution of the data was checked for normality by means of the Shapiro-Wilk test, with a significance level of 0.05. Correlations between different variables were explored in order to investigate the relationship between the frequency of multimedia use, and between subsections of the questionnaire, and the overall evaluation of the program.

TABLE 3 Performance of the models for Dutch

Phonological class	Accuracy training set	Accuracy validation set	Precision	Recall	F-score
Nasal	0.945	0.935	0.957	0.935	0.941
Plosive	0.942	0.910	0.936	0.910	0.917
Fricative	0.926	0.936	0.949	0.936	0.940
Approximant	0.916	0.889	0.965	0.889	0.917
Trill	0.899	0.937	0.969	0.937	0.949
Labial	0.941	0.917	0.948	0.917	0.927
Coronal	0.871	0.884	0.895	0.884	0.887
Dorsal	0.885	0.902	0.934	0.902	0.912
Glottal	0.947	0.979	0.991	0.979	0.984
Voice	0.911	0.935	0.935	0.935	0.935
Vowel	0.898	0.892	0.912	0.892	0.896
Bilabial	0.956	0.916	0.958	0.916	0.930
Labiodental	0.954	0.975	0.985	0.975	0.978
Alveolar	0.874	0.885	0.898	0.885	0.888
Postalveolar	0.975	0.994	0.998	0.994	0.996
Pause	0.975	0.978	0.979	0.978	0.978

Usability and operability

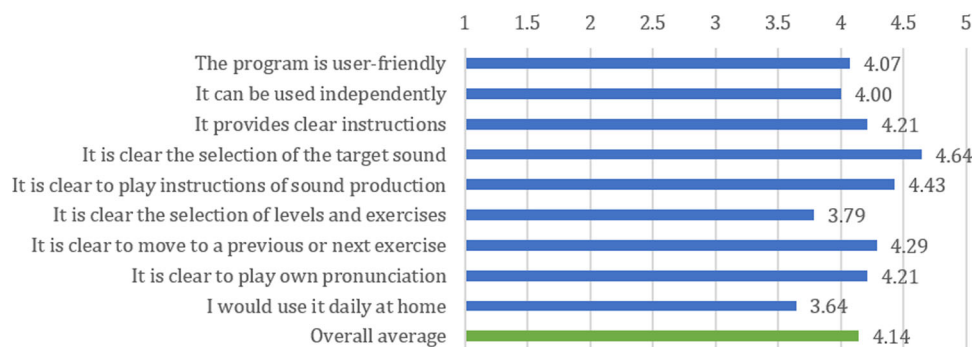


FIGURE 5 Usability and operability scored on a Likert scale (1 = 'completely disagree' and 5 = 'completely agree'). [Colour figure can be viewed at wileyonlinelibrary.com]

RESULTS

Algorithm performance

The trained Dutch models achieved high accuracy rates for the detection of all trained phonological classes. This statistical approach of the human perception process has detection rates > 88% (*F*-score). During the learning process frequently occurring speech sounds of the language were modelled with higher accuracy than infrequently occurring sounds, but in general the results are considered accurate enough for the discrimination between each specific phonological class. Table 3 summarizes the performance of the models for the training dataset and the validation dataset (also called the development set). Pre-

cision, recall and *F*-score performance measures are also offered.

Survey

Usability and operability

Although some participants needed help with the mouse control, an average score of 4.14 out of 5 was given to the section 'Usability and operability' (figure 5). This means that the patient's perception of the software functionality was generally found as easy to use and understand. Most participants agreed that the VAT is user-friendly and can be used independently, but the answers to

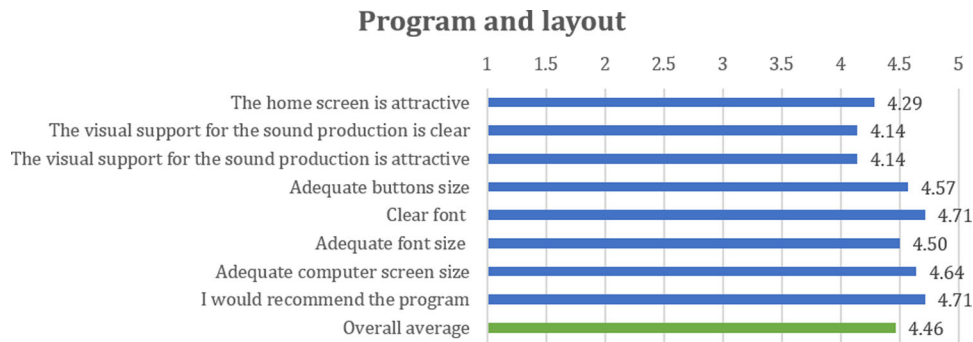


FIGURE 6 Program and layout scored on a Likert scale (1 = 'completely disagree' and 5 = 'completely agree'). [Colour figure can be viewed at wileyonlinelibrary.com]



FIGURE 7 Exercises scored on a Likert scale (1 = 'completely disagree' and 5 = 'completely agree'). [Colour figure can be viewed at wileyonlinelibrary.com]

the statement 'I would use the program daily at home' were quite diverse. Half the subjects agreed or totally agreed, five were neutral and two disagreed. This could be related to the experience and daily use of multimedia, which was scored the lowest on the questionnaire, results which are presented below and that will be analysed later in this study. Figure 4 summarizes the average score of the items in this section.

Program and layout

The section 'Program and layout' received an average score of 4.46 out of 5 (figure 6). Most of the users found the program attractive, with a clear layout. The visual and auditory support for the target sound production was perceived as appealing and supportive. Finally, all subjects would recommend the use of the VAT to other patients.

Exercises

The section 'Exercises' received an average score of 4.23 out of 5 (figure 7). The users found the set of exercises to be clear and useful. This is particularly important because this type of repetitive and intensive training tends to be tiresome; therefore, it is essential that users find out the usefulness of the selected set of exercises that can help to improve their articulatory ability.

Feedback

The majority of users perceived the visual feedback given during therapy as clear and attractive. The section 'Feedback' received an average score of 3.96 out of 5 (figure 8). However, for this section the patients provided several suggestions to improve the stimulation and motivation of the user. They recommended also adding auditory feedback by the VAT in case of correct target production.

Experience in the use of multimedia

The section related to the participants' use of multimedia received an average score of 3.06 out of 5 (figure 9), being the lowest of the questionnaire. These results clarify the score given before to the statement of daily usage of the software at home. Although the majority of the participants have some kind of multimedia device for entertainment, they do not use it every day. This is an interesting point that should be considered for the assignment of therapy frequency and duration.

A moderate correlation was found between experience in the use of multimedia and the perceived quality of the program $r = 0.554$ ($p = 0.04$), which was measured as the total score of the questionnaire. Patients with higher knowledge or experience with the usage of multimedia tended to have a more positive perception of the quality of the software-based therapy. However, patients with less

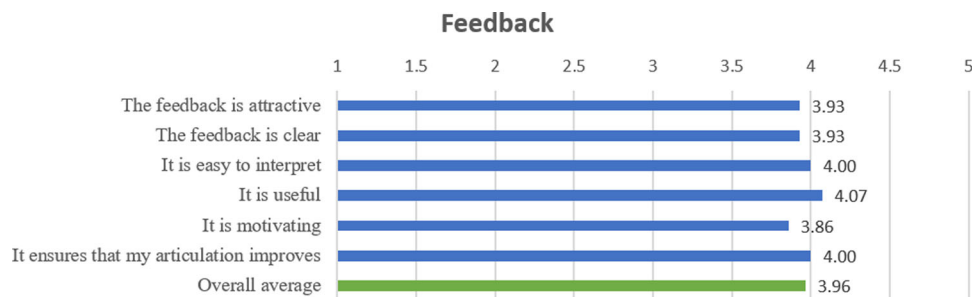


FIGURE 8 Feedback scored on a Likert scale (1 = 'completely disagree' and 5 = 'completely agree'). [Colour figure can be viewed at wileyonlinelibrary.com]

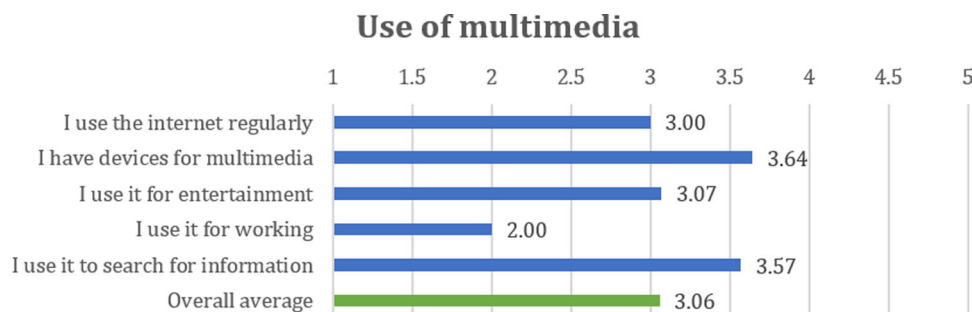


FIGURE 9 Use of multimedia scored on a Likert scale (1 = 'completely disagree' and 5 = 'completely agree'). [Colour figure can be viewed at wileyonlinelibrary.com]

experience or interest in multimedia also showed a positive attitude to the therapy.

Sections with a higher impact on the patient's perception of the program

From a practical and statistical point of view, most of the aspects that were addressed in the questionnaire are considered highly significant for the end-users' perception of the quality, acceptability and usability of the computer-based therapy program. The biggest impact on the overall score was related to the feedback, followed by the exercises; also usability and operability were highly significant for end-users. The program and layout seem to be important for practical use, but not statistically significant in the perception of acceptability and usability of the VAT (table 4).

DISCUSSION

This pilot study evaluated the feasibility of a VAT that guides patients with dysarthria through a BARt program, combining speech technology, phoneme-specific feedback and motor learning principles. It is generally accepted

that for motor speech disorders, greater therapy frequency leads to better ultimate performance (Dobkin and Thompson 2000). The VAT addresses the need for an accessible program granting patients the opportunity to train their articulatory skills more intensively and frequently. The tool provides a set of exercises, visual and auditory modelling of target utterances, and adequate feedback with accurate performance. It has an accessible interface for patients with brain damage who often deal with other motor and sensory dysfunctions and have limitations in language and cognition. Additionally, it can considerably increase the capacity of speech therapy for patients with practical limitations (distance, financial restrictions, etc.).

The results of the pilot study show that the VAT addresses the elementary aspects necessary to deliver the BARt successfully and to ensure the user's engagement. The study-specific questionnaire addressed the functionality, layout, content of the speech material used and feedback given during training. Quantitative evaluation after two sessions of interaction with the system shows that this study cohort, consisting of patients with mild to moderate dysarthria, but with sufficient cognitive, language, visual and auditory skills, were in general positive about this new computer-based therapy approach. These results were consistent with previously published virtual therapies in the

TABLE 4 Pearson correlation coefficients and significance level between the total score and the subscales of the questionnaire

	Use and operation	Program and layout	Exercises	Feedback
Total score	0.672** ($p = 0.009$)	0.408 ($p = 0.147$)	0.761** ($p = 0.002$)	0.785** ($p = 0.001$)

Note: Correlation is significant at the *0.05 level (two-tailed); and **0.01 level (two-tailed).

field of speech and language therapy (Cole *et al.* 2007, Saz *et al.* 2009, Engwall *et al.* 2006, Eriksson *et al.* 2005). The majority of participants ensured that the VAT is attractive, user-friendly and can be used independently. The usefulness of the software was evaluated by two different statements within the survey, whether they found the type of exercises to be useful as well as the given automatic feedback. Patients' responses confirmed that it is a useful tool that can help to improve their communication. The visual feedback was found to be attractive, clear and easy to interpret; however, there are some suggestions for its improvement and for keeping the user motivated, which is a key factor for successful training sessions. The recommendations include the addition of auditory feedback with a more stimulating and encouraging response, as a suggestion, applause that is played when the target sound is correctly pronounced, which can make the success in target production more evident to the users. Previous research indicates that rewarding feedback after accurate expressions is crucial to maintain a high motivation (Halpern *et al.* 2012, Öster 2006). Participants also pointed out that providing an overall score of articulation will grant a better view of their progress.

The section related to the use of multimedia in daily life shows that this specific group of participants (mainly elderly patients or with brain damage) are not highly motivated to use multimedia devices; however, most of them indicated a positive attitude toward the VAT program. This could be interpreted as they feel the program can be helpful to improve their articulatory skills while practising in their environment.

The general outcomes from the questionnaire can also be essential for traditional face-to-face therapy methods of articulation. SLPs may use these outcomes to provide better, more efficient and individualized articulatory therapy, taking into account that a well-structured and hierarchical programme of exercises is necessary and needs to be goal oriented, and the most valuable information they can offer to the patients is good and encouraging feedback.

The involvement of patients in the early stage of the design and developmental process of the system is a way to guarantee the usability and acceptability of the program. The recommendations/suggestions made by the participants will lead to a suitable, easy-to-use and engaging virtual therapy.

LIMITATIONS

The study was restricted in time to two VAT therapy sessions interacting with the software program in a rather small cohort. This implies that the results of this pilot study are preliminary and cannot be generalized. In future validation studies, a larger number of patients should be included. Adding qualitative interview data and perspectives of caregivers and patients' family members may also lead to additional insights. Finally, the questionnaire can be upgraded by adding an item related to how accurate the subjects felt the feedback was.

FUTURE RESEARCH

Future research will focus on the evaluation of therapeutic effectiveness after introducing the VAT in daily patient care. This will include a higher number of subjects and objective measures of longitudinal articulatory improvements. Also, future research will focus on the optimization of the acoustic models for the automatic error detection in dysarthric speech, the development of an objective acoustic measure of sound distortion, and the correlation analysis of automatic versus perceptual measurements of the degree of distortion. Qualitative evaluation of the system by a wider community that includes caregivers, SLPs and patients' family members will also be investigated. The integration of exercises with higher linguistic levels will also be the focus of future work. The VAT will be released as an open-source system, with a general public licence. It will also be communicated via professional organizations of speech therapists and patient groups.

CONCLUSIONS

The computer-based articulation therapy for adults with dysarthria seems to be feasible and well accepted by the end-users of this pilot study. Most participants have a positive attitude towards the use of the VAT. Even participants with limited computer skills/experiences are able to use the VAT independently. High accuracy was obtained on the performance of the models used for error detection, which contributes to adequate automatic feedback. However, the survey points out that improvements to the system are still



necessary, and all the suggestions given by the stakeholders will be considered.

The inclusion of patients in the design and development of these types of software is significantly important for usability, acceptability and engagement with e-therapy, which appears to be a good option to increase the frequency and intensity of practice while supporting the traditional methods. The VAT offers clear applicability with an added value for the Flemish health facilities, including a positive impact for a specific group of patients.

ACKNOWLEDGEMENTS

This research was funded by the European Union's Horizon 2020 Research and Innovation Programme under Marie Skłodowska-Curie (grant agreement number 766287). We express our gratitude to the speech and language therapist Margot Baudelet for her valuable contribution to the visual modelling of the target sounds. We also thank all the participants and the speech and language therapist for their commitment and participation.

DECLARATION OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ORCID

Viviana Mendoza Ramos  <https://orcid.org/0000-0002-0533-091X>

Juan C. Vasquez-Correa  <https://orcid.org/0000-0003-4946-9232>

REFERENCES

- Abad, A., Pompili, A., Costa, A., Trancoso, I., Fonseca, J., Leal, G., et al. (2013) Automatic word naming recognition for an on-line aphasia treatment system. *Computer Speech & Language*, 27, 1235–1248.
- Amaya, A., Woolf, C., Devane, N., Galliers, J., Talbot, R., Wilson, S., et al. (2018) Receiving aphasia intervention in a virtual environment: the participants' perspective. *Aphasiology*, 32(5), 538–558.
- Bhagal, S. K., Teasell, R. & Speechley, M. (2003) Intensity of aphasia therapy, impact on recovery. *Stroke; A Journal of Cerebral Circulation*, 34(4), 987–993.
- Chaisanit, S., Suksakulchai, S. & Nimnual, R., (2010) Interactive multimedia courseware of vowel training for the hearing impaired. In: 2010 International Conference on Control Automation and Systems (ICCAS), pp. 1196–1199.
- Cherney, L.R., Halper, A.S., Holland, A.L. & Cole, R. (2008) Computerized script training for aphasia: preliminary results. *American Journal of Speech-Language Pathology*, 17, 19–34.
- Cole, R., Halpern, A., Ramig, L., Vuuren, S.V., Ngampatipatpong, N. & Yan, J. (2007) A virtual speech therapist for individuals with Parkinson's Disease. *Journal of Educational Technology*, 47, 51–55.
- Darley, F.L., Aronson, A.E. & Brown, J.R. (1969) Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research*, 12, 246–269.
- Darley, F.L., Aronson, A.E. & Brown, J.R. (1975) *Motor speech disorders*. W. B. Saunders.
- De Bodt, M.S., Huici, M.E.H.D. & Van De Heyning, P.H. (2002) Intelligibility as a linear combination of dimensions in dysarthric speech. *Journal of Communication Disorders*, 35(3), 283–292.
- De Bodt, M., Guns, C. & Van Nuffelen, G., (2006) *NSVO: Nederlandstalig O. Spraakverstaanbaarheidsonderzoek*. Vlaamse Vereniging voor Logopedisten.
- Deruyter, S. & Staessens, N. (2018) Exploratief onderzoek naar de bruikbaarheid en toepasbaarheid van ASISTO voor spraakrevalidatie bij personen met dysartrie na CVA.
- Dickson, S., Barbour, R.S., Brady, M., Clark, A.M. & Paton, G., (2008) Patients' experiences of disruptions associated with post-stroke dysarthria. *International Journal of Language & Communication Disorders*, 43(2), 135–153.
- Dobkin, B.H. & Thompson, A.J. (2000) Principles of neurological rehabilitation. In: Badley, W.G. et al. (Eds.) *Neurology in clinical practice: principles of diagnosis and management*, vol 1, ed 3, Boston: Butterworth-Heinemann.
- Duffy, J.R. (2019) *Motor speech disorders E-book: Substrates, differential diagnosis, and management*. Elsevier Health Sciences.
- Dykstra, A.D., Hakel, M.E. & Adams, S.G. (2007) Application of the ICF in reduced speech intelligibility in dysarthria. In: *Seminars in speech and language* Thieme Medical Publishers, Vol., 28, No. 04, pp. 301–311.
- Engwall, O., Bälter, O., Öster, A.-M. & Kjellström, H. (2006) Designing the user interface of the computer-based speech training system ARTUR based on early user tests. *Behaviour & Information Technology*, 25, 353–365.
- Eriksson, E., Bälter, O., Engwall, O., Öster, A.-M. & Kjellström, H.S. (2005) Design recommendations for a computer-based speech training system based on end user interviews. In: Proceedings of the Tenth International Conference on Speech and Computers, pp. 483–486.
- Fox, C.M., Ramig, L.O., Ciucci, M.R., Sapir, S., McFarland, D.H. & Farley, B.G. (2006) The science and practice of LSVT/LOUD: neural plasticity-principled approach to treating individuals with Parkinson disease and other neurological disorders. *Seminars in Speech and Language*, 27(04), 283–299.
- Garvey, M.A., Giannetti, M.L., Alter, K.E. & Lum, P.S. (2007) Cerebral palsy: new approaches to therapy. *Current Neurology and Neuroscience Reports*, 7(2), 147–155.
- Halpern, A.E., Ramig, L.O., Matos, C.E., Petska-Cable, J.A., Spielman, J.L., Pogoda, J.M., et al. (2012) Innovative technology for the assisted delivery of intensive voice treatment (LSVT® LOUD) for Parkinson disease. *American Journal of Speech-Language Pathology*, 21, 254–267.
- Huybrechts, G., Decoster, W., Goeleven, A., Lembrechts, D., Manders, E. & Zink, I. (1999) *Articulatie in de praktijk: vocalen en diftongen*. (Acco).
- Kalinyak-Flizar, M., Martin, N., Keshner, E., Rudnicki, A., Shi, J. & Teodoro, G. (2015) Using virtual technology to promote functional communication in aphasia: preliminary evidence from interactive dialogues with human and virtual clinicians. *American Journal of Speech-Language Pathology*, 24, 974–989.
- Kent, R.D. Ed. (1992) *Intelligibility in speech disorders: Theory, measurement and management* (Vol., 1). John Benjamins Publishing.
- Kirakowski, J. & Corbett, M. (1993) SUMI: the software usability measurement inventory. *British Journal of Educational Technology*, 24(3), 210–212.

- Kleim, J.A. & Jones, T.A. (2008) Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research*.
- Kleim, J.A., Jones, T.A. & Schallert, T. (2003) Motor enrichment and the induction of plasticity before or after brain injury. *Neurochemical Research*, 28(11), 1757–1769.
- Klinkenberg, S., Bredeweg, B. & Molenaar, P. (2004) *Constructie en validatie van een algemene computervaardigheid vragenlijst*. ACV.
- Lee, J.B., Kaye, R.C. & Cherney, L.R. (2009) Conversational script performance in adults with non-fluent aphasia: treatment intensity and aphasia severity. *Aphasiology*, 23(7–8), 885–897.
- Lee, S.A.S. (2019) Virtual speech–language therapy for individuals with communication disorders: current evidence, limitations, and benefits. *Current Developmental Disorders Reports*, 6(3), 119–125.
- Maas, E., Robin, D.A., Hula, S.N.A., Freedman, S.E., Wulf, G., Ballard, K.J., et al. (2008) Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech–Language Pathology*, 17(3), 277–298.
- Marshall, J., Booth, T., Devane, N., Galliers, J., Greenwood, H., Hilari, K., et al. (2016) Evaluating the benefits of aphasia intervention delivered in virtual reality: results of a quasi-randomised study. *PLoS One*, 11(8), e0160381.
- Massaro, D.W. & Light, J. (2004) Using visible speech to train perception and production of speech for individuals with hearing loss. *Journal of Speech, Language, and Hearing Research*, 47, 304–320.
- Mendoza Ramos, V., Paulyn, C., Van Den Steen, L., Hernandez-Diaz Huici, M.E., De Bodt, M., & Van Nuffelen, G. (2021). Effect of boost articulation therapy (BArT) on intelligibility in adults with dysarthria. *International Journal of Language & Communication Disorders*, 56(2), 271–282.
- Miller, G.F., Vásquez-Correa, J.C. & Nöth, E. (2020) Assessing the dysarthria level of Parkinson's disease patients with gMM-UBM supervectors using phonological posteriors and diachokinetic exercises. In: *International Conference on Text, Speech, and Dialogue* (pp. 356–365). Cham: Springer.
- Oostdijk, N. & Broeder, D. (2003) The Spoken Dutch Corpus and its exploitation environment. In: Proceedings of 4th International Workshop on Linguistically Interpreted Corpora (LINC-03) at EACL 2003.
- Orozco-Arroyave, J.R., Vásquez-Correa, J.C. & Nöth, E. (2020) Current methods and new trends in signal processing and pattern recognition for the automatic assessment of motor impairments: the case of Parkinson's disease. *Neurological Disorders and Imaging Physics*, 5.
- Orozco-Arroyave, J.R., Vásquez-Correa, J.C., Klumpp, P., Pérez-Toro, P.A., Escobar-Grisales, D., Roth, N., et al. (2020) Apkinson: the smartphone application for telemonitoring Parkinson's patients through speech, gait and hands movement. *Neurodegenerative Disease Management*, 10(3), 137–157.
- Öster, A. M. (2006) Computer-based speech therapy using visual feedback with focus on children with profound hearing impairments. Doctoral dissertation, Tal musik och hörsel.
- Palmer, R., Enderby, P. & Hawley, M. (2007) Addressing the needs of speakers with longstanding dysarthria: computerized and traditional therapy compared. *International Journal of Language and Communication Disorders*, 42, 61–79.
- Renkema, J. ED. (1995) *Woordenlijst Nederlandse taal: Samengest. door het Instituut voor Nederlandse Lexicologie in opdracht van de Nederlandse Taalunie*. Sdu Uitgeverij.
- Rosenbek, J.C. & Jones, H.N. (2009) Principles of treatment for sensorimotor speech disorders. *Clinical Management of Sensorimotor Speech Disorders*, 2, 269–288.
- Saz, O., Yin, S.-C., Lleida, E., Rose, R., Vaquero, C. & Rodríguez, W.R. (2009) Tools and technologies for computer-aided speech and language therapy. *Speech Commun*, 51, 948–967.
- Da Silva, M.P., Junior, A.A.C., Balen, S.A. & Bevilacqua, M.C. (2012) Software use in the (re)habilitation of hearing impaired children. *Jornal da Sociedade Brasileira de Fonoaudiologia*, 24, 34–41.
- Thompson, C.K., Choy, J.J., Holland, A. & Cole, R. (2010) Sentactics®: computer-automated treatment of underlying forms. *Aphasiology*, 24(10), 1242–1266.
- Tjaden, K. (2007) Segmental articulation in motor speech disorders. In Weismer, G. (Ed.) *Motor speech disorders: essays for Ray Kent*. Plural Publishing, 151–186.
- Tykalova, T., Ruz, J., Klempir, J., Cmejla, R. & Ruzicka, E. (2017) Distinct patterns of imprecise consonant articulation among Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. *Brain and Language*, 165, 1–9.
- Van Den Steen, L., Van Nuffelen, G., Guns, C., De Groote, M., Pinson, L. & De Bodt, M. (2011) De spraak handicap index: een instrument voor zelfevaluatie bij dysarthriepatiënten. *Logopedie*, 24, 26–30.
- Van Vuuren, S. & Cherney, L. (2014) A virtual therapist for speech and language therapy. In: Bickmore, T., Marsella, S., Sidner, C. (Eds.), *Intelligent Virtual Agents*: Springer International Publishing, pp. 438–448.
- Vásquez-Correa, J.C., Klumpp, P., Orozco-Arroyave, J.R. & Nöth, E. (2019) Phonet: a tool based on gated recurrent neural networks to extract phonological posteriors from speech. In: Proc. Interspeech pp. 549–553.
- Verhoeven, J. (2005) Belgian standard Dutch. *Journal of the International Phonetic Association*, 35(2), 243–247.
- Weismer, G. (2007) *Motor speech disorders: Essays for Ray Kent*. Plural Publishing.
- Weismer, G., Jeng, J.Y., Laures, J.S., Kent, R.D. & Kent, J.F. (2001) Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica*, 53(1), 1–18.
- Yorkston, K.M. & Beukelman, D.R. (1978) A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of Communication Disorders*, 11(6), 499–512.

How to cite this article: Mendoza Ramos V, Vasquez-Correa JC, Cremers R, et al. Automatic boost articulation therapy in adults with dysarthria: Acceptability, usability, and user interaction. *International Journal of Language & Communication Disorders*. 2021;56:892–906. <https://doi.org/10.1111/1460-6984.12647>



APPENDIX A

TABLE A1 Applied questionnaire

<p>Use and operation</p> <ul style="list-style-type: none"> The program is user-friendly The program can be used independently The program provides clear instructions It is clear how to select a target sound It is clear how to play, pause, and stop the instructions about the target sound production It is clear how to select levels and exercises It is clear how to go to a previous or next exercise It is clear how you can listen to your pronunciation I would use the program daily at home 	<p>Program and layout</p> <ul style="list-style-type: none"> The home screen is attractive The visual support for target sound production is clear The visual support for target sound production is attractive The size of the buttons is adequate/sufficient The font is clear The font size is adequate/sufficient The computer screen size is adequate I would recommend the program
<p>Exercises</p> <ul style="list-style-type: none"> The exercises are clear The type of exercise is useful The number of exercises per target sound is sufficient The degree of difficulty of the exercises matches my ability 	<p>Feedback</p> <ul style="list-style-type: none"> The feedback is attractive The feedback is clear The feedback is easy to interpret The feedback is useful The feedback is motivating The feedback ensures that my articulation improves
<p>Use of Multimedia</p> <ul style="list-style-type: none"> I use the internet regularly I have one or more devices for multimedia (computer, smartphone, tablet, etc.) I use multimedia for relaxation/entertainment I use multimedia for work I use multimedia to search for things/information 	

Comments and suggestions: