Professional challenges in computer-assisted speech therapy

Doru-Vlad Popovici, Cristian Buică-Belciu*

*University of Bucharest, FPSE, Șos. Pandurilor, nr. 90, S5, Bucharest 050.663, Romania

Abstract

While earlier software used in speech therapy was usually limited to preset exercise patterns chosen by speech language pathologists according to their professional judgment, the latest intelligent systems cover all main speech therapy issues, including diagnosis, therapy exercises, performance monitoring, instant feedback both to the client and the speech language pathologist. Some of the professional challenges practitioners are facing today are related to the proper use of technology in their everyday practice, including the use of communication devices to deliver speech and language therapy services at a distance.

Keywords: computer-assisted speech therapy; automatic speech recognition; speech language pathology; logopaedics; telepractice.

1. Introduction

Early diagnosis of children's speech and language disorders is of prime importance for any seasoned SLP practitioner (or logopaedist), especially when access to speech and language therapy services is difficult because of adverse circumstances such as commuting, travel expenses, lack of time, poor family education, low family income, restrictive social customs, or, plainly, lack of qualified personnel in some school districts or at schools located in rural areas. As normal development of speech involves the simultaneous development of both the peripheral segment (phono-articulatory organs) and the central segment (the cortical areas responsible for auditory-verbal training schemes and verbal-motor complexes), imperfections found in the speech of children under age of three are considered by many SLP practitioners to reflect the general dynamics of this linguistic maturation process. Excluding pervasive developmental disorders (such as autism), congenital defects (e.g., cleft lip and palate), or other documented medical conditions, delays in proper speech acquirement are generally ignored by unsuspicious parents and

* Corresponding author. Tel.: +4-031-425-3445; fax: +4-031-425-3445.
E-mail address: buicabelciucristian@yahoo.com.
kindergarten teachers, especially when the child is under age of three or is going through dentition changes. Moreover, “the waiting game” is maintained by those adverse circumstances mentioned above. Doing so, corrective-remedial interventions require more and more time and effort as the presentation at the speech therapy practice office is continually postponed. So to speak, it is just one good example of “the boiling frog” syndrome. Although there are grounds for restraint in formulating a diagnosis at an early age, hoping for a spontaneous remission of speech difficulties may lead to poor communication skills in the future, having negative impact in the psychosocial and professional adjustment area, not to mention disadvantages regarding proper development of self-image and personality (Buică, 2004).

Speech therapy can be time consuming, especially when phono-articulatory dynamics are complex. In most cases, the phono-articulatory model is the one provided by speech and language pathologists themselves (for pronunciation exercises conducted in the practice room) or other significant individuals, namely the parents (for exercises performed at home). Constant monotony of exercises and lack of opportunities for individual practice affect children's motivation for recovery and diminish their ability to focus on therapy goals. Use of intelligent diagnosis and therapy systems is seen today as a viable alternative to traditional approaches in order to increase speech and language therapy efficiency (Popovici, Buică, & Velican, 2010).

2. From ASR to CASLT: A review of intelligent systems used in speech language therapy

Creating a program capable to identify not only isolated phonemes, but also the co-articulated ones (in syllables, monosyllables, and words), led to the emergence of automatic speech recognition (ASR) systems. Their essential task was to react appropriately to the verbal input of any speaker, through dedicated interfaces (Voice User Interfaces). The first program of this kind appeared in 1952, being able to identify isolated given digits (Davies, Biddulph, & Balashek, 1952). Therapeutic applications of ASR for people with disabilities are primarily those based on the effect of feedback for improving pronunciation of dysarthric patients (Ferrier, Jarrell, Carpenter, & Shane, 1992; Thomas-Stonell, Kotler, Leeper, & Doyle, 1998) and real-time transliteration of speech into print for deaf and hearing impaired individuals (Stuckless, 1994), but the motivational boost produced on subjects during correcting activities cannot be ignored (Parsons, 1997). Also, ASR systems can make a phonetic vocabulary speech transcription of people with speech-language disorders, which can be compared to the standard patterns stored in a database (Griffin, Wilson, & Clark, 2000). Other applications of ASR technology are accent reduction and, respectively, correction of inadequate pronunciation of people whose native language is different than the language of the country of residence. One such program is Pronto, originally created for American English speakers wishing to learn Spanish, and for Chinese speakers (Mandarin dialect) studying American English (Dalby & Kewley-Port, 1999).

Automatic speech recognition programs still seem to be considered merely ways of converting oral language into verbal/writing language using specialized software (Kitzing, Maier, & Lyberg, 2009). Current commercial speech recognition programs are mostly designed for adults’ pronunciation phonemes discrimination, which do not make them useful for children's verbal input (Gustafson & Sjölander, 2002). The most important factor is the need to increase the reliability of speech recognizers before they can fully be included into the daily lives of people with disabilities (Noyes, Haigh, & Starr, 1989). Today, automatic speech recognition technology is a powerful tool in mainstreaming students with special education needs (see more in Revuelta, Jimenez, Sanchez, & Ruiz, 2011).

In time, intelligent systems for diagnosis and therapy of speech-language disorders have appeared, known as acronyms like CBST (Computer Based Speech Training), CAMST (Computer-Assisted Methods for Speech Therapy), CAST (Computer-Aided/Assisted Speech Therapy), or CASLT (Computer-Aided Speech and Language Therapy). A multimedia program created in 1985, continuously
updated, completed with tailored versions for international languages, with facilities for use in speech therapy at home is LingWare/STACH (Griessl & Stachowiak, 1994). This program was designed as an additional form of support in ordinary speech therapy and not as an easy replacement. CATSEAR is an integrated interface, designed for signal collection, data analysis, therapeutic design and monitor speech therapy (Turk & Arslan, 2005). Applications are limited to moderate correction of pronunciation difficulties, and voice training. PEAKS (Program for Analysis and Evaluation of all Kinds of Speech Disorders) requires a standard PC provided with internet connection and sound card. The subject is put through a standardized test which is then automatically ranked. The system can be used for pronunciation and voice disorders. The program can be used by a speech pathologist as means of alternative assessment, particularly to identify those associated problems which cannot be diagnosed from the beginning. The program has proved especially useful in working with patients who underwent total laryngectomy (consecutive laryngeal cancer) and children with rhinolalia (Maier, Haderlein, Eysholdt, Rosanowski, Batliner, Schuster, et al., 2009). The Telelogos system not only allows speech pathology, but lay persons (clients, family members, assistive personnel, etc.) to adapt existing practice programs to the needs of each individual case using either a configuration editor or a vocabulary editor. Additionally, the Telelogos system includes a set of tests accessible only to special education professionals, designed to diagnose a wide range of language disorders, and learning disabilities as well. (Glykas & Chytas, 2005a; Glykas & Chytas, 2005b). Last but not least, The Logomon system was designed to support speech therapy, both at the logopaedic cabinet and at home throughout the duration of therapy (Danubianu, Pentiuc, Schipor, Nestor, & Ungureanu, 2008). The database contains over a thousand exercises for assisting, on one hand, general therapy (motility development, control of the breathing rhythm, stimulating phonemic awareness) and, on the other hand, the specific therapy (phoneme building, consolidation, and automation) (Pentiuc, Tobolcea, Schipor, Danubianu, & Schipor, 2010).

The involvement of the technological factor in speech therapy is becoming more and more accentuated, improving the quality of treatment, the efficiency of the management of time spent in the speech therapy office, and a faster access to corrective exercises that the client must practice. The contribution of computer-assisted speech therapy software is particularly noted for the success obtained in activities such as phonemic insertion, phonemic practice, phonemic correction, and phonemic automation, and less in the diagnosis of the language disorders (Popovici, Buică, Velican, & Cârțușă, 2010; Popovici, Buică, & Velican, 2011).

3. Use of technology in speech language therapy: advantages and limitations

According to Saz, Yin, Lleida, Rose, Vaquero, & Rodriguez (2009), there are three general areas of diagnosis and intervention for speech and language disorders: the acquisition of basic phonatory skills (control of breathing, voicing, speech intensity, and tone), the acquisition of the phonetic system of the language (pronunciation, creation of syllables and words), and the language understanding. The computer programs used in speech therapy are generally focused on providing the exercises the client must complete, while the evaluation of pronunciation correctness is still done by the speech therapist. The applications come with a substantial library of exercises and interesting graphic interfaces meant to draw the attention and concentration of the subject. The advantage is indisputable: easiness and rapidity in use for the therapist and the subject, as well as the possibility to practice at home. The major disadvantage consists in the impossibility of the program to determine the type and degree of language disorder and in the severe limitation of monitoring the therapy progress and automated readjustment of practice exercises upon the phonological feedback provided by the subject. In other words, the results of the whole therapy depend essentially on the therapist’s direct expertise, including variables such as therapist’s agenda, his or her professional experience, and effective intervention time for each client, therapy group size, session’s
frequency and so on. According to a recent study (Gillam & Frome Loeb, 2010), four language therapy intervention components proven to be critical: intensity, active attention, feedback, and rewards. All these components can be stimulated in both the speech therapy office and at home, using a class of algorithms which, implemented into a software, would dramatically diminish downtime in therapy and would minimize disruptive influences of external variables, individual (fatigue, disinterest, haste, etc.), social (negative interpersonal interactions), environmental, organizational and so on.

Usually, speech language pathologists don’t have advanced computer skills, their professional area of expertise being restricted to communication disorders. However, older and younger generations of SLPs face the challenge of adapting IT technologies to their therapeutic purposes although these new tools were never conceived for speech language pathologists (Bratti, 2010). Psychological limitations of using CASLT technology stem, on one hand, from “cyberophobia” (that is, fear of computers), and on the other hand, from the trap of “computerized factotum”. While easiness, rapidity, interactivity, mobility, and flexibility are the usual keywords employed to stress the advantages of applying modern technological advances to traditional speech and language therapy, the enticement of overconfidence in technological reliability (sometimes exaggerated for commercial purposes by developers themselves) and the illusion of a “silver bullet” technological panacea could demote the professional status of the speech therapist and lead to harmful (or at least inefficient) self-administered therapies.

4. Conclusions

It should be noted that such software should be used as an assistive tool and not as a substitute for the speech language pathologist, even considering telepractice services. Ideally, the role of this kind of program is to relieve both SLP practitioners and clients of repetitive tasks, well defined in terms of intervention parameters and requiring routine activities. In this situation, specialists are in charge of configuring the software and monitoring the evolution of cases from time to time, devoting the rest of their time taking care of the remaining atypical or “rebellious” cases. Such a software application can also be used by the client at home for longer periods of time compared with existing programs, reducing downtime between two successive check-ups at the speech therapy office or diminishing the limitations of remote speech therapy (telepractice). The relevance of computer-assisted speech therapy depends, on the one hand, on the increasing capability of software applications to identify defective articulation samples by implementing increasingly sophisticated signal processing algorithms, and, on the other hand, on the ability of speech language pathologists to insert such applications skillfully into their therapy strategies, individualized for each client. The key factor of the use of intelligent systems for diagnosis and therapy of language disorders is that the software does not set the priorities of speech therapy and does not direct the practice activities, but the speech language pathologist is the one who decides for each case at a time the moment, quota, content and duration of automated interventions, according to speech therapy principles and the specific structure of the projected speech therapy program.

Acknowledgement

This work was supported under the exploratory research project “Intelligent system for assessing and correcting speech disabilities used in remedial therapy” funded by the National Council of Scientific Research in Higher Education, Romania (ID_1683, Contract number CNCSIS nr. 846/2008).

References


