

Multimedia Speech Therapy Tools and Other Disability Solutions as Part of a Digital Ecosystem Framework

David Calder

► **To cite this version:**

David Calder. Multimedia Speech Therapy Tools and Other Disability Solutions as Part of a Digital Ecosystem Framework. 6th IFIP TC 12 International Conference on Intelligent Information Processing (IIP), Oct 2010, Manchester, United Kingdom. pp.326-335, 10.1007/978-3-642-16327-2_39. hal-01055055

HAL Id: hal-01055055

<https://hal.inria.fr/hal-01055055>

Submitted on 11 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Multimedia Speech therapy Tools And Other Disability Solutions as Part of A Digital Ecosystem Framework

David Calder,
Curtin University of Technology,
Bentley, Perth,
West Australia
david.calder@cbs.curtin.edu.au
Tel. 61-8-9266 2875

Abstract. Curtin University has developed a multimedia prototype system for use by speech therapists to assist in the rehabilitation of motor speech impaired patients. These are often stroke victims who have to relearn the ability to communicate effectively. The process is usually laborious and instead of the usual card-based prompts used by the therapist, this multimedia solution offers screen and audio prompts together with high quality digitised speech. The result is a reduced work load for the therapist. In parallel with this work, is a Mobility Aid Assistive Device Program for the visually impaired. The support clusters for all these groups, whether they be therapists, caregivers, manufacturers of hardware or software do not interact to the extent they might. The author proposes a collaborative digital ecosystem framework that would assist this challenge.

Keywords: Multimedia, computer-based therapy, speech therapy, cardiovascular accident, assistive technology, sound interface displays, laser, disabled, infrared, long cane, portable electronic device, sensory channels, visually impaired, ultrasonic pulse-echo, ambient sound cues.

1 Introduction

There is a worldwide shortage of speech training services to accommodate the clients who require speech therapy [1]. Cardiovascular accidents are the third largest killer in Australia after heart disease and cancer. Those that survive this trauma usually have need for rehabilitation. Motor impairment of the speech organ is a common occurrence. We have produced a Multimedia system which replaces the chart and paper-based cues found in many speech therapy units. The aim is to assist speech therapists by decreasing the time they have to spend in direct contact with patients. More importantly, there are also a number of benefits to clients and these will be discussed later.

As part of the Digital Ecosystem and Business Intelligence Institute (DEBII) research team at Curtin University, we worked closely with speech professionals from Royal Perth Hospital in order to tailor the interface exactly to their requirements. We were guided by practicing speech therapists in ensuring existing and well established interactive therapy strategies for certain classes of dysarthria, were followed as closely as possible. The first version of the system has undergone beta testing in Western Australia. This is a straightforward software solution using standard personal computers already available to therapists within their environment.

The author has also been researching assistive devices for the visually impaired and portable speech aids for the speech impaired. These, on the other hand, require dedicated hardware and the design of *extreme demand* interfaces in order to offer the most appropriate solution. However, the support infrastructure clusters, for each of these above examples, do not collaborate or communicate to any great degree.

2 Cerebro-vascular accident

When the oxygen supply to the brain is blocked or when an artery in the brain is ruptured, a common stroke results. The former is the result of a clot traveling to the brain and occluding blood flow in a cerebral artery [2]. Once the oxygen supply has ceased, hydrogen ions propagate within the brain cells and damage the cells. Within four to eight minutes, the cells begin to die, and of course this destruction is irreversible. In the case of a reduced blood flow, cells may stop functioning but later recover when normal perfusion is restored [3].

Cardiovascular accidents are the third largest killer in Australia and in many other countries of the western world. The highest killer is heart disease followed by cancer [4]. Fifteen percent of people over the age of forty living in the western world die of cardiovascular accidents. A further fifteen percent are also stroke victims who survive and require some form of rehabilitation [5]. Of this group five percent make a full recovery whilst ten percent suffer some permanent disability [6].

As medical technologies have improved and the relative death rate has fallen, so the need to offer suitable rehabilitation services to the victims who have survived this trauma has increased. Patients are left with multiple disabilities and these often include loss of speech. The anterior portion of the brain is associated with speech production and damage to this area can result in motor speech disorders [7]. The severity of the motor speech disorder may vary from person to person but even where spontaneous recovery is evident, some speech rehabilitation will probably be carried out. In the case of more severe speech impairment, long term rehabilitation may result [8]. In most instances, the road to recovery is long and difficult, placing stress on both client and therapist. Any means that may assist this load for both the client and the therapist should be investigated.

3 Hardcopy patient cue cards

During a standard speech therapy session, the client is presented with a number of cues which include symbols representing phonemes/vowels, drawings indicating air flow through the mouth and nose and the many combinations of lip, teeth and tongue positions which are essential in general articulate speech. Cards were identified as part of a major problem by speech professionals during the initial data analysis of the old system at Royal Perth Hospital, Shenton Park in Western Australia. The therapist is under pressure to find the correct flash card or phoneme chart as she sits with the client. Together with these paper-based cues she uses her own speech and face to offer sound and visual cues to the client. In order to produce the consonant M for example, the client may be asked to place the lips together and allow the air to flow through the nasal cavity. The therapist will give an example by producing the sound herself and simultaneously accentuate the positioning of the lips.

The client is then asked to repeat this step. Consonant and consonant/vowel combinations such as this need constant repetition by the client and consequently the therapist before any improvement is expected. This face-to-face contact can be very tiring for both the client and the therapist. As therapy is done on a one to one basis,

duration of contact per individual per week may be limited to only a few hours. Clients are therefore being deprived of the continuous training which would benefit their rehabilitation.

4 Multimedia Therapy Tool

The concept of a system to assist in speech therapy resulted from earlier work with speech therapists in the United Kingdom in developing a speech aid prototype [8]. The development of the therapy tool depended on a requirement to address some of the most significant problems associated with conventional therapy methods. The Department of Speech Pathology at the Royal Perth (Rehabilitation) Hospital, Australia, was involved in the development of the system and provided the old model of operations for analysis from which the computer-based system was developed. Assistant Professor G. Mann, now in the Department of Communication Disorders at the University of Florida and widely recognised as a key authority, worked closely with us over several years in developing the original version of the multimedia tool [10].

The therapy tool aims to relieve constant therapist/patient supervision, particularly where time consuming repetitive tasks are involved. Therapists can use their time more effectively in planning new goals whilst the computer provides visual and sound cues to the client. Therapists no longer have to organise cue cards or sort through hundreds of icons and drawings. These were seldom in colour whereas the computer-based system augments these traditional methods by using colour and animation. The latter was something that could not be achieved on loose pieces of paper or cardboard! Consequently the therapy process can run more smoothly and effectively as all cues are presented on the screen and/or produced by the high quality stored speech system.

There are a number of other benefits which particularly relate to patients and have been highlighted during beta testing. It has been found that the system could be used at home as well as in the conventional environment of the speech therapy unit. Where a client is left with the computer and removed from the clinical surroundings, the stress of embarrassment is removed. This is particularly evident when an older client has to "relearn how to speak" in the presence and under the direction of a young speech therapist. The stress placed on the client during these encounters should be seriously considered.

Another bonus for the system is that it is based on a standard IBM compatible PC and could even be run from a portable notebook mounted to a wheelchair. Most households now have several computers with a more than adequate built-in sound systems [11]. This means that in most instances, only the software need be installed in the home to allow therapy to continue. Other members of the family could be involved in the rehabilitation and the clinical sessions at the hospital could be used for the monitoring of progress and not be associated with intrusion into the privacy of the patient [12]. Communication within the family is an important part of rehabilitation and the therapy tool could help in promoting this. The high cost of therapy is relieved but not replaced by this augmentative system [13].

Since the therapy tool saves clinician time, it may also help relieve the shortage of speech therapy services. If the amount of time spent with each client is reduced, the clinician could then take on a greater case load.

The nature of a computer-based system is that, "it doesn't have to see another patient in an hours' time". Therefore clients may train at their own pace under no stress. Sound and visual cues may be repeated over and over again without the pressure of supervision. Of course not all patients would be suitable candidates for this level of freedom but initial tests have indicated great success for certain motor impaired victims.

5 Multimedia Animation and Stored Speech

Asymetrix Multimedia Toolbook was used to develop this system. It operates in a high level, Windows-based object-oriented programming environment. Toolbook uses a book metaphor for multi screen production. Each screen is described as a page and all pages in a production are called a book. Making use of these properties pages were constructed to emulate the icons and visual cues used in conventional therapy.

Animated visual representations of the vocal system were added to the standard set of cues. High quality digitised speech was recorded from a practicing speech therapist so that an exact representation of a set of speech sounds could be achieved. As the normal therapy session involves a progressive set of prompts to initiate speech from the client, the therapy tool was set up with a progressive set of cue buttons which gradually give more and more assistance to the client in achieving a particular goal. For example, this might be help in pronouncing the consonant F. A set of six cue buttons mounted vertically on the right hand side of the screen allow for this positive reinforcement. These cues may take visual or spoken form. Another reason for a progressive interactive strategy is the problem of complexity which has plagued so many previous interface designs for the disabled [14].

The sound cues may be both instructive and exemplary. One cue button will produce an instruction such as, "place the lower lip against your upper front teeth and blow" whilst another produces the target as required, the actual sound "fff". Another button links the consonant to a word so as to place the target sound in context. Yet another offers the word in the context of a sentence.

All cue buttons and command buttons have explanatory logos which appear at the bottom of the screen when the cursor moves over their area. There are over 300 digitised sound recordings associated with the therapy tool.

The system is normally controlled by a mouse and test models incorporate this means of user control. Other remote switch devices are currently being investigated. This will make the system more versatile and therefore suitable for quadriplegics [15]. A system of this type should be flexible enough to adapt to suit the needs of each individual client and varying physical disabilities [16]. The system being used by a speech therapy patient and therapist can be seen in Figure 1 below.

Fig. 1. Speech therapy session with stroke patient.

Research by the author and therapists has established that phoneme synthesis is not suitable for the speech requirements of this system. A robotic sounding voice may be acceptable for certain games programs and simple communication tools but can never be adequate for the variety of accents, languages and key sounds that may be required by therapists. The Institution of Electrical and Electronics Engineers [17], published a recommended practice for speech quality measurements, but this has never been fully adopted as the basis for standardised testing [18]. Edwards showed that no existing rule-based synthesiser came close to passing the test of being indistinguishable from natural speech.

To date, the system has been tested on a small pilot sample of patients in Australia. Initial tests at the Royal Perth Hospital have indicated that the response to the system for middle age stroke victims was very positive. Other age groups including children are included in forward planning for the later version. The interface has proved easy to understand and selected stroke patients have managed to control the system themselves, i.e. choosing and activating cue buttons.

Future tests will be coordinated to fine tune the system before full trials take place. It must be emphasized that full and robust trials have yet to be carried out. This is a work in progress. However, feedback from the first pilot tests highlighted the need for fast adaptability in design refinement which could be partly addressed by the digital ecosystem paradigm [19]. Issues such as universal access are also relevant [20].

6 Other Support Clusters

Speech therapy tools form just part of a wider picture of related hardware or software based support systems, all falling roughly within the catchment area of assistive technologies for the disabled. Most of these systems require special interface design and may cater for temporary or permanent loss of speech. The software tool described above is assisting in a hopefully temporary loss of speech. However, this loss may also be permanent. In which case alternative and augmentative devices such as portable speech aids (producing synthetically generated speech) are required. Even though these tools are closely related and familiar to most speech therapists, there is a huge gulf between the teams that produce and support both solutions.

This communication gulf turns into a chasm when more diverse assistive technologies are examined. These may include sensory loss such as hearing or visual impairment. The author has been working on both the above projects involving

speech therapy and a substitution aid for natural speech. More recently we have been developing a navigation device for the blind. And this is where the infrastructure support clusters become very different.

There are numerous mobility aids and orientation mapping devices for the visually impaired on the market at present, some with significant drawbacks. Many assistive technology devices use ultrasonic pulse-echo techniques to gauge subject to object distance. Some use infrared light transceivers or laser technology to locate and warn of obstacles. These devices exhibit a number of problems, the most significant of which are related to the interface display that conveys navigation/obstacle warning information to the user. Other sensory channels should not be compromised by the device. This is exactly what can happen when, for example, audio signals are used in obstacle warning on/off displays.

The DEBII team has developed a prototype device, which it is hoped, will be the first step in addressing some of the above listed problems. Patent searches indicate this working prototype has a unique tactile interface design which, even in its basic form, should have distinct user advantages. As with some of the sonar systems listed above in the paper, this first prototype is best suited to outdoor use. Future models are not limited to sonar technology, however.

The design criteria has and will in the future, concentrate on intuitive interfaces that do not compromise certain other all-important sensory channels. These interfaces, now under development, will be configurable for users who are both deaf and visually impaired. There will also be an emphasis on ease of learning and use. It is unacceptable to expect someone, who may have multiple disabilities, to undertake a long and complex learning program in how to use a new device. The Innovation in the author's mobility aid design may be summarized as follows:

'A portable mobility aid incorporating warning obstacle ahead information with mapping capabilities as a stand-alone device.'

At this stage, no further technical specification can be given due to IP novelty protection. It is hoped that in future papers, we will be able to concentrate more freely on the detailed technical aspects of the design. However, the current system uses ultrasound for range-finding and is mounted on a cane for test purposes. Initial tests have proved the system to be at least as effective as many of the alternative commercial systems available.

7 Digital Ecosystem Models

Issues of complexity with respect to individual requirements must be seen within the context of a wider ecology of the particular user, with that person clearly at the centre, contributing to a team solution. An established and highly successful ecological approach to designing individualized education programs for the disabled student has been refined over twenty years into a highly recommended model and is now regarded as 'best practice' [21].

This ecological approach has not as yet permeated all areas of disability support. However, the power of the digital ecosystem framework is now accepted within many other disciplines, particularly with respect to small enterprise collaboration [22].

Within small business, the advent of the web has allowed sales penetration over vast distances. Accompanying these advances have come new modes of marketing and partnership possibilities that would have been impossible only a few years ago.

With this connectivity has come a fertile and dynamic business theatre that cannot be avoided if small enterprises are to survive. This interaction has led to collaborative workflow models [23].

The logic behind collaborative workflows is to produce a sequence of activities that not only produce a meaningful result, but also to facilitate small groups working together to achieve common goals. The actual physical distance and associated limitations between these entities then becomes less important as web based tools are used to link enterprises and their common aspirations [24]. The entities themselves may be small companies competing against large predator corporations, or widely dispersed cottage industries (such as those associated with assistive devices) with a common interest [25].

Beyond the standard empowerment the digital ecosystem model has provided, are more specific areas that are pertinent to such groups operating in harmony. One of the most important of these is trust evaluation [26]. Other typical support areas are logistics and privacy [27, 28]. These would act as foundations for the framework I that is proposed.

Digital Ecosystems For Assistive Technology clusters (DEFAT) is a proposed collaborative cluster-based ecosystem model, neither limited by distance between clusters nor the particular disability types associated with each of the clusters. Individual clusters may include a range of specialist personnel associated with the support of a client requirement. The output of such an environment would not only be the efficient research and development of appropriate assistive devices, but also result in more streamlining for the teams in their everyday support of an individual, be that speech therapy for dysarthria patients or training in the use of a long cane or mobility aid for the visually impaired.

8 DEFAT Structure

With each client representing a nucleus at the centre of his or her support cluster, an individual's local ecological environment has been acknowledged (as discussed and cited in previous sections) as a worthwhile starting point, offering a framework from which specialist support action may be fleshed out.

Each support cluster would have a number of clients within its particular category of disability. Cluster membership would not be determined by distance or physical boundaries. The aim would be to maximize use of the digital ecosystem paradigm in order to break existing physical boundaries. By applying a DEFAT strategy, current digital technologies such as mobile, the internet and video conferencing can be coordinated and optimized to deliver the best outcome for all members of this ecosystem.

Open-ended but novel design solutions would be encouraged from both hardware and software developers. The sharing and exchange of common modular solutions at both a functional and user interface level would be part of the ecosystem membership requirement.

The difference would be in the focus and modular consideration of appropriate novel and relevant ideas, when first considering I.P. matters. This will not always be relevant to designs, but when it is, it should in fact enhance the potential for success

within the DEFAT community itself, as well as in a wider context (external to the ecosystem).

Those academic cluster members who currently work within a limited research environment with a very small interest group would have the opportunity to share their research and ongoing projects on a wider stage within the digital ecosystem. Cross-disciplinary interaction would be nurtured by DEFAT.

A cluster of people with a vast range of interdisciplinary skills would focus on a user group of people all with a common disability. There would be many separate clusters, meeting the challenges of specific needs of different disability groups. As now, it may be assumed that special education specialists, therapists, medics, academics, engineers and particularly hardware and software experts would form part of each cluster, the main difference being a recognition of the greater ecosystem in which each cluster coexists and operates.

Users at the center of each cluster, the nucleus, would determine the nature of the environment. Clusters would communicate with each other for a common good and the ecosystem itself would be able to benefit from its size in terms of external links and its critical mass.

A starting point for such a structure may take into account the problem as defined by Liu et al when referring to building the right systems and the need for better tools and environments in their paper on component-based medical and assistive devices and systems [29]. They put forward a ten-year roadmap, which fits well as a start to implementing the DEFAT paradigm.

Clusters need to be client centered, taking into account breakthrough research such as that of Bach-Y-Rita into sensory substitution [30] and Merzenich into brain plasticity [31].

A global advantage and DEFAT's greater mass would benefit the ecology on many levels. There would be lower manufacturing costs than is now associated with small-run dedicated systems production. This advantage would result from greater demand for DEFAT modular units across clusters and existing boundaries. Relatively large production runs catering for a global DEFAT module demand would drive production costs down.

9 Conclusion

The Digital Ecosystem and Business Intelligence Institute at Curtin University in Western Australia is developing a multimedia-based training tool for use by speech therapists and their clients. A more advanced system is currently under development. This new version is specifically aimed at people with motor speech disorders. It is believed that the design may, in time, be developed into a wider range of solutions which may incorporate related therapy requirements.

The author compares the development of the above speech therapy and speech aid solutions with other assistive device development examples; specifically those with sensory impairments, such as blindness. As most of these software and dedicated hardware solutions are produced by small, unlisted companies, there is little in the way of publicly available, reliable sales figures, and as such the addressable market success is not well defined. However, interviews conducted with industry experts, in addition to the small size of the companies themselves, suggest that these competing

devices have so far failed to achieve any significant market presence, and in many cases, have inherent and significant user interface design issues. The author's prototype programme and proposed DEFAT framework development will, hopefully, be a step in the right direction.

References

1. Editor, 2008, May 22. ASU Designs Program To Address Shortage of Speech Therapists, Phoenix Arizona News. DOI= <http://evliving.com>
2. Hewer, R.L. & Wade D.T. 1986. Stroke - A Practical Guide Towards Recovery. Methuen Australia : Australia
3. Kaplan N.M. & Stamler M.D. 1983 Prevention of Coronary Heart Disease. W.B. Saunders and Co.
4. Gross, P.F. 1991 The Total Economic Costs of Stroke in Australia. Technology Assessment Corporation Pty Ltd. NSW, Australia.
5. Hird, K.M. 1987 Communicative Competence, Self-care and Mobility in Left and Right Hemisphere Adult Stroke Patients. Curtin University of Technology, Western Australia
6. Dyken, M.L., Woff P.A., Barnett J.H. 1984. Risk Factors in Stroke: A Statement for Physicians by the subcommittee on Risk Factors and Strokes. Stroke (6): 1105-1111
7. Darley, F.I., Aronson A.E, Brown J.R 1975. Motor Speech Disorders. W.B. Saunders Co. : London
8. Gloag, D. 1985 Rehabilitation After Stroke: What is the potential? British Medical Journal 290, 699-701
9. Calder, D.J. , Lister P.F., Mack Watson, J.D 1987 Large Vocabulary Speech Aid Incorporating Softkey Switching. Third Symposium International De Ingeniera Biomedica. Madrid. 755 – 781
10. Mann. G. 2002. The Mann Assessment of Swallowing Ability. Cengage Delmar Learning, New York.
11. Edwards A., Pitt I. 2002. Design of Speech-Based Devices. Practitioner's Series. Springer Professional Computing, London
12. Alkalay, L & Asserman M.W. 1983. Cerebrovascular Accidents: Care and Rehabilitation. The Practitioner 227. 469-473.
13. Hodge S. 2007. "Why is the potential of augmentative and alternative communication not being realized? Exploring the experiences of people who use communication aids". Disability and Society, 22, 466
14. Seidle, F. 1977. Is Home Care Less Expensive? Health Social Work 2(5) -10.
15. Vanderheiden, G.C and Lloyd, L.L. 1986. Communication Systems and Their Components. Augmentative and Alternative Communication: American Speech and Language-Hearing Association. Rockville(Maryland) 49 , 162
16. Reichle, J.J. York, J., Sigafoos J 1991 Implementing Augmentative and Alternative Communication. Paul Brookes Publishing Co. London 239 - 256.
17. IEEE 1969. IEEE Recommended Practice For Speech Quality Measurements. IEEE Transactions on Audio and Electronics, AU,17 (3), 225- 246.
18. Edwards, A.D. 1991. Speech Synthesis: Technology for Disabled People. Paul Chapman Publishing Ltd: London 12- 40.
19. E. Chang, M. West. 2006. "Digital Ecosystems and Comparison to Collaboration Environment". WSEAS Transactions on Environment and development 2, 1396-1404.
20. Zimmermann, G., Vanderheiden, G., Gilman, A. 2002. Prototype Implementations for a Universal Remote Console Specification. CHI 2002 Conference on Human Factors in Computing Systems, Minneapolis, MN, pp510 – 511.
21. Rainforth, B. York J., Macdonald C., *Collaborative Teams for Students With Severe Disabilities*, Baltimore: Paul Brookes, 1993, pp. 71- 83.
22. E. Chang, M. West. "Digital Ecosystems and Comparison to Collaboration Environment". *WSEAS Transactions on Environment and development 2*, 2006, pp1396-1404
23. Pudhota L., Chang E, "Modelling the Dynamic Relationships between Workflow Components" ICEISI Porto, Portugal, 2004

- 24.D. Neumann," An Introduction to Web Objects". On-line at [http://mactech.com/articles/mactech/Vol.13/13.05/WebObjects Overview](http://mactech.com/articles/mactech/Vol.13/13.05/WebObjects%20Overview). 2004
- 25.Ulieru M. Brennan Scott R., "The Holonic enterprise: a model for Internet-enabled Global Manufacturing Supply Chain and workflow Management", Canada, 2000
- 26.E. Chang, T. Dillon and F Hussain, "Trust and Reputation for service-oriented Environments: Technology for building Business intelligence and Consumer Confidence," John Wiley and Sons, West Sussex, England, 2006
- 27.M. Clark, P. Fletcher et al. Web Services Business Strategies and Architectures. Expert press, 2002.
- 28.Skinner G. and Chang E.. "A Projection of the Future Effects of Quantum Computation on Information Privacy and Information Security" *International Journal of Computer Science and Network Security* 6, 2006, pp166-172"
- 29.J. Liu, B. Wang, H. Liao, C. Shih, T. Kuo, A. Pang and C. Huang, "Component-based Medical and Assistive Devices and Systems," in *Proceedings of High Confidence Medical and Systems (HCMDSS) Workshop*, 2005, Philadelphia, PA
- 30.Bach-Y-Rita P and Kerrel S., "Sensory Substitution and the Human-Machine Interface," *Trends in Cognitive Sciences*, vol. 7, pp. 541-546, 2003
- 31.Merzenich MJenkins., W., *Memory Concepts*, Elsevier, 1993, pp 437 - 453