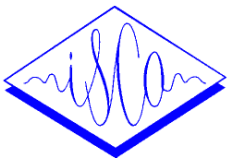
ISCA Archive
http://www.isca-speech.org/archive



ESCA/SOCRATES Workshop on Method and Tool Innovations for Speech Science Education (MATISSE)
University College, London, UK
April 16–17, 1999

InterPHACE - Internetworked links for Phonetic Analysis in Clinical Education

Janet Mackenzie Beck, Alistair Lawson and Ineke Mennen

Department of Speech and Language Sciences, Queen Margaret University College, Edinburgh, Scotland
j.beck@sls.qmced.ac.uk, a.lawson@sls.qmced.ac.uk, i.mennen@sls.qmced.ac.uk

Abstract

This collaborative project addresses two key issues. The first is the need for Speech and Language Therapy (SALT) students to develop skills in analysis of a wide range of speech disorders and to apply these to clinical practice. The second is the lack of availability of instrumental analysis equipment in many SALT clinics.

Telematic links between Queen Margaret University College (QMUC) and a variety of Health Care Trust clinics are being established, allowing speech data to be transmitted from SALT clinics to QMUC. Speech data is analysed in the QMUC Speech Analysis Laboratory, and the results are accessible by SALTs. QMUC staff and students are then able to engage in teleconference discussions with the SALT responsible for each client's management, thus learning about the impact of analysis results on diagnosis and management. A data-base of clinical speech material is also being developed as a teaching/learning resource.

1. Introduction

Accurate assessment and monitoring of speech production using phonetic analysis of speech output is a prerequisite for diagnosis and appropriate remediation of speech disorder. Ensuring a high level of expertise in phonetic analysis is therefore of major importance, both in undergraduate education of SALTs and in continuing professional development (CPD). There is a need for novel and shared resources for education in phonetic analysis techniques at both undergraduate and post-qualification levels which foster the ability to apply phonetic knowledge and skills in a way which informs

The use of technology to complement auditory-perceptual analysis of speech poses a particular challenge for undergraduate and continuing education. Research increasingly validates clinical application of new forms of instrumental analysis, providing evidence that auditory-perceptual analysis does not always give sufficient information for accurate diagnosis and optimal targeting of remedial strategies. For example, recent research on covert contrasts has shown that detailed instrumental analysis of speech may reveal differences in phoneme realisation which are not perceived even by highly skilled listeners, but which have important implications for differential diagnosis and management of speech disorder [1]. Unfortunately, very few SALTs have access either to instrumental analysis systems or to the time and/or expertise required for interpretation of results and, as a consequence, undergraduate SALT students may have little opportunity to apply speech analysis skills within clinic placements.

Figure 1 is a diagrammatic representation of the way in which telematic links between QMUC and Health Care Trust clinics are being developed with the aim of meeting the needs of both SALTs and students. These can provide a speech analysis service to SALTs, while allowing QMUC students to be actively involved in collection and analysis of speech data and teleconferenced discussions with SALTs about management implications. The data base of clinical speech material and analysis results will be further developed as a resource for teaching, learning and research at undergraduate and post-qualification level, thus encouraging future use of instrumental analysis to

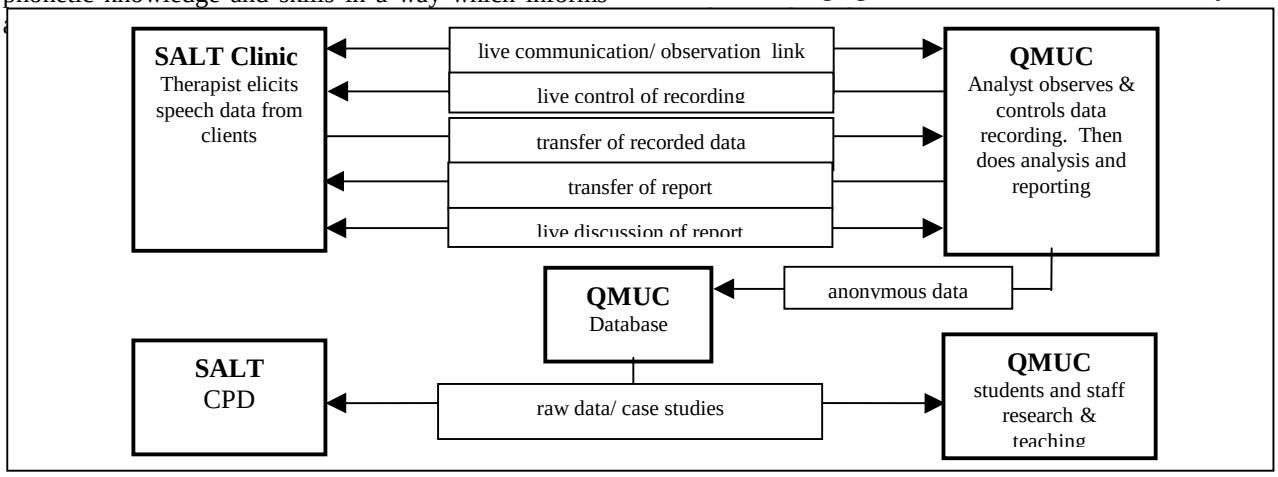


Figure 1. Telematic links for Phonetic Analysis in Clinical Education

Careful evaluation of all aspects of the system is integral to the project. A variety of evaluation strategies will be used which draw on the principles proposed by the LTDI Programme [2] and involves a suite of integrative evaluation techniques (observation, questionnaires, interviews and focus group discussions) which has been modified specifically for this purpose.

2. The InterPHACE System

This part of the paper describes the technological aspects of the project. First the general requirements for the InterPHACE system will be outlined followed by the main technical constraints. Thirdly the evolutionary approach used to specify and develop the system to meet these general requirements and elicit any further particular requirements will be discussed. Described finally is the initial system set up in the community SALT clinic based at the QMUC Scottish Centre for Research into Speech Disability (SCRSD), where communication strategies and the technology can be scrutinised and modified for each of the client groups involved in the project before the system is installed in the other Health Care Trust Clinics involved in the project.

In order to clarify the discussion, the person at the clinic end of the InterPHACE system is called 'therapist' and the person at the other end 'speech analyst'. Either or both could in fact be speech and language therapists, students or speech scientists

2.1 General InterPHACE System Requirements

The discussion of the system requirements of the project are divided first into requirements for speech data recording, second into the requirements for data analysis and reporting, and third into requirements for Case study or Raw data access for Teaching/ CPD and Research. These requirements were distilled from discussions with Speech Scientists and Therapists at QMUC and Therapists from the SALT clinics that will be involved in the project. Further requirements will be elicited as the project progresses.

2.1.1 Requirements for speech data recording:

Within SALT Clinic:

- to record CD quality sound (or data from other instrumentation e.g. EPG, EMA, EGG)
- equipment that can be unobtrusive for certain clients
- but that could display tokens or give immediate feedback if required for other clients
- equipment that requires little or no input from the Therapist/Student
- but that allows the Therapist to take over more control of the recording and analysis if wished

At the same time within QMUC, the speech analyst requires to :

- remotely control the recording equipment
- monitor the quality of the recording
- communicate with the therapist e.g. if microphone isn't switched on (this communication may on occasion be without knowledge of the client)
- control the process of backing up the data as soon as possible after recording

2.1.2 Requirements for data analysis and reporting:

- a one week turn-around time from data collection to submitting the report back to therapist and discussing it (clients usually have weekly appointments and the therapist may want to adjust therapy in light of the report)
- secure transfer of data from SALT Clinic to QMUC
- the speech analyst must be able to identify each target speech sound (e.g. 'peak' and 'beak' may sound the same). Additional information about which portion of speech data corresponds to which target may have to be communicated from therapist to speech analyst
- data analysis and presentation on standard software packages such as Cool Edit (Version 96, Syntrillium Software Corp.), Multi-Speech (Model3700, version 1, Kay Elemetrics Corp.), Dr Speech (Tiger DRS Inc), and Excel (Version 7.0, Microsoft Corp.)
- automation of data analysis and presentation procedures that do not require human judgement and intervention
- teleconference discussion of report between QMUC and SALT Clinic (not necessarily video conference - audio and computer screen are likely to be sufficient)
- data and reports to be stored on a database

2.1.3 Requirements for case study or raw data access for teaching/ CPD and research:

- anonymous, secure raw data, details of how it was analysed, and copy of report
- accessible for compilation of case studies, research projects and CPD material by certain SALT clinic staff, QMUC students and staff
- accessible for viewing case studies, research projects and CPD material by SALT clinic staff, QMUC students and staff

2.2 Technical Constraints

The main technical constraint is lack of access to high speed data links between QMUC and the Edinburgh SALT clinics. The present budget of the project prohibits installation of high speed links and the 'politics' of getting them installed in SALT clinics would probably drag on beyond the timescale of the project. However if the current project is successful it may provide the precedent and required leverage to obtain access to high speed data links for the project.

Our aim is to develop a system which is accessible to any SALT clinic with a standard telephone line and PC whilst being able to make use of high speed data links should they become available. This means that the

design of teleconference links based on audio and textual or computer graphic conferencing is more appropriate than video conferencing which requires high speed data connections. Real-time recording and transmission of CD quality sound would again require very high speed data links whereas a standard modern PC can record CD quality sound directly to its hard disk and then later transfer the data down a telephone line via a modem. Although it may take around 10 minutes to send 1 minute of CD quality sound, this is not a problem as the SALT clinics are looking for the report back within a few days. The data transfer could be done automatically overnight.

2.3 System Development

An evolutionary approach has been chosen to develop a system that will meet the general requirements and constraints listed above and that will be flexible enough to meet the particular requirements of each client group. We have initially decided to adopt an approach which is as 'low tech' as possible, using mass produced standard components (hardware and software) wherever possible. Customised software or hardware will be introduced when necessary to integrate standard components and to relieve bottle necks e.g. in speech data processing. One of our aims is to ride on the wave of rapid technological progress of standard desktop PC hardware, software and standard telephone communications technology. This technology includes transfer of data files between PCs, remote control of one PC by another, application sharing (e.g. both PCs viewing and able to edit the same spreadsheet), text, audio and video conferencing.

The first four stages involved in our evolutionary approach are listed below. These concentrate on the data acquisition, analysis and reporting aspects of the project. This is where most of the technical challenge and innovation will occur. These stages involve:

1. setting up initial system which allows us to carry out PC-based speech recording in SCRSD by remote control from another room in QMUC, whilst allowing observation of the communication/ interaction required during and between each process involved
2. demonstrating the system and concepts involved (including remote control of PCs, teleconferencing, application sharing, data transfer, etc) to therapists and analysts involved in the project
3. using the system for clinical case studies
4. reviewing, and evolving system and repeat stages above. Installing system in remote clinics when it is robust enough to be set up for particular therapists and client groups

2.4 Initial system set-up

Figure 2 shows the initial set up in SCRSD SALT clinic and QMUC as used in the case study described below. The greyed-out portions of the figure show additional equipment which will enable us to study the communication and interaction between therapist and analyst whilst using the system. Remote control of speech recording is carried out over standard telephone lines using LapLink Technical remote access software (version 1.5, Travelling Software Inc.) which also allows text and voice 'chat' when not recording.

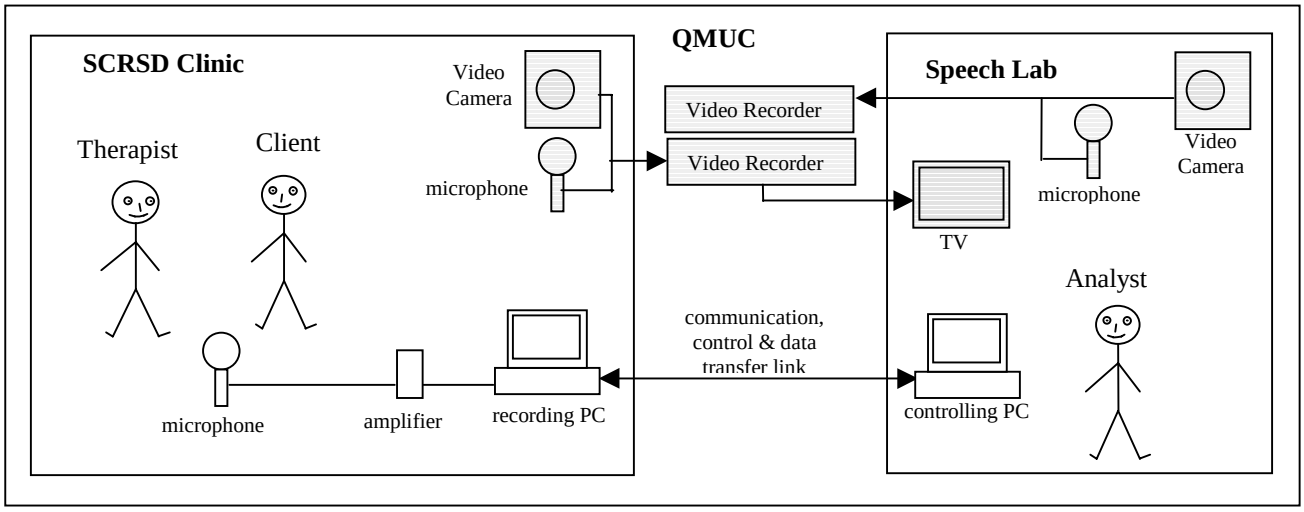


Figure 2. Initial InterPHACE system set-up

3. Case Study

A single case study may be helpful as an illustration of the way in which the recording may be incorporated into a clinic session, and to show how results of acoustic analysis may be made available for discussion with the therapist. The client is a six year old boy (K) with severe phonological disorder, who is receiving therapy from a community therapist working within the SCRSD. In consultation with the therapist it was decided to do a limited study on voicing contrast. A picture naming task was used to elicit monosyllabic target words with word initial alveolar and bilabial stops or /s/+stop clusters, all before the vowel /i/. Six repetitions of each of six tokens were elicited during the first 10 minutes of one of K's regular therapy sessions. Correct target identification for each token was ensured by the SALT reading the number of each token aloud, and by use of a record sheet to note any problems with elicitation or changes in intended sequence. This part of the session, which took place in a quiet therapy room, was recorded digitally at 44100 Hz sampling rate and 16 bit resolution directly onto a PC hard-disk through a stand-microphone. The session was also video-taped. The observer, who sat in the control room, monitored and controlled the actual recording program. After the session tokens were extracted for further measurement using Cool-Edit, and voice onset time (VOT) was measured using Multi-Speech and collated and presented using Excel. A protocol sheet for feedback of results for bilabials is shown in Figure 3. Perceptual judgments of voicing in this child's output had indicated that voicing contrasts are not yet established, but VOT measurements suggest that, at least for bilabial plosives, a covert contrast (i.e. one which is not yet perceptible) may be emerging. This may be taken as evidence of meta-phonological awareness of the contrast between /p/ and /b/. Discussion between the SALT and QMUC staff therefore focused on the implications this might have for management.

This was one of a number of pilot case studies which demonstrated that remote control of the recording process caused minimal disruption of the clinical session, whilst providing the therapist with clinically relevant analysis results and offering both students and therapists an opportunity to see the application of acoustic analysis techniques to clinical management. Two types of case have been identified where the system may be of greatest benefit; those where a detailed analysis of one aspect of speech production may allow a more focused approach to therapy and monitoring (e.g. where the presence or absence of covert contrasts may indicate a level of phonological processing, as in the example above) and those with very complex patterns of communication problem where a wide range of analyses may be required.

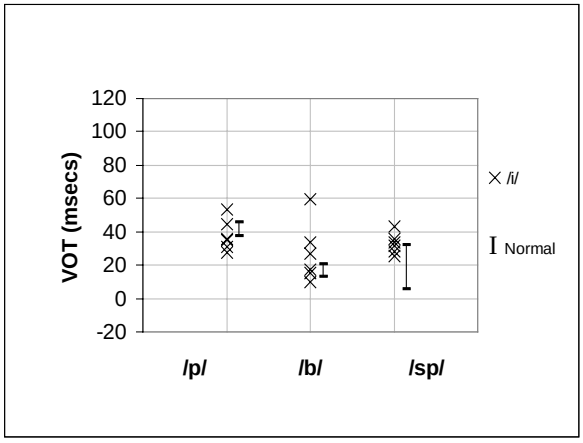


Figure 3. VOT for word initial plosives in the words "peak", "beak", & "speak".

4. Conclusions

The system reported here aims to apply new, but readily available, technology to support clinical practice, continuing professional development and clinical education. The potential of such a system is considerable, and it is envisaged that telematic links would be expanded to involve more distant parts of Britain (e.g. the Western Isles) and beyond. There might be considerable commercial potential in offering a central speech analysis service to Health Care Trusts, whilst a wider range of clinical applications could include remote assessment /monitoring of clients in non-specialist clinics (e.g. screening in health centres) or delivery of individually tailored computer-based therapy packages to house-bound clients. Educational applications, using the database to support delivery of teaching/learning modules, might be of particular value to countries with newly-developing SALT professions.

Acknowledgements

We would like to thank BT for funding this project through its Higher Education Development fund, and all NHS Trust SALTs who have supported this project. In particular we thank Carolyn Hawkes and her client for help in initial piloting of the system.

References

[1] Scobbie, JM, Gibbon, F, Hardcastle, WJ & Fletcher, P (1999). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe and J. Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon*, Cambridge University Press, Cambridge. (in press)

[2] Stoner, G (Ed) (1996) *Implementing Learning Technology.* Learning Technology Dissemination Initiative: Heriot-Watt University, Edinburgh.