

Kent Academic Repository

Full text document (pdf)

Citation for published version

Sharp, Paul and Kelly, Stephen W. and Main, Alison and Manley, Graham (1999) An instrument for the multiparameter assessment of speech. *Medical Engineering & Physics*, 21 (9). pp. 661-671. ISSN 1350-4533.

DOI

[https://doi.org/10.1016/S1350-4533\(99\)00091-0](https://doi.org/10.1016/S1350-4533(99)00091-0)

Link to record in KAR

<https://kar.kent.ac.uk/16611/>

Document Version

UNSPECIFIED

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

AN INSTRUMENT FOR THE MULTIPARAMETER ASSESSMENT OF SPEECH

Paul Sharp*, Steve Kelly*, Alison Maint† and Graham Manley‡.

*Medical Electronics Research Group, Electronic Engineering Laboratory, University of Kent, Canterbury, Kent, CT2 7NT, UK; † Speech and Language Therapy Department, William Harvey Hospital, Ashford, Kent, TN24 0LZ, UK; ‡ Dental Department, Canterbury Health Centre, 26 Old Dover Road, Canterbury, Kent, CT1 3JH, UK.

ABSTRACT

This paper describes the development of SNORS+, which is a clinical, user-friendly, instrument, for measurement of the articulators, during speech. The design criteria for the instrument were based upon a wide-ranging review of current practice and available techniques. SNORS+ allows objective assessment of the function and co-ordination of key articulators. Appropriate targeting of therapy is therefore possible. Visual feedback is provided, for therapy and an objective measurement of outcome is easily obtained. Preliminary results are presented. These suggest that the instrument will prove extremely useful in the assessment and management of many speech disorders.

Keywords: SNORS+, instrumentation, speech disorders, multiparameter

INTRODUCTION

SPEECH PRODUCTION

Speech is the result of a highly complex and versatile system of co-ordinated muscular movements¹. Speech sounds comprise periodic and/or aperiodic sound waves. Periodic sound waves are produced by the vocal folds, which lie in the larynx. This signal is then modified by the articulators. Aperiodic sound waves are produced by allowing air to pass through the open vocal folds into the vocal tract to be modified by the articulators, creating friction. Additional complex sounds are produced by combining the two methods (see Figure 1).

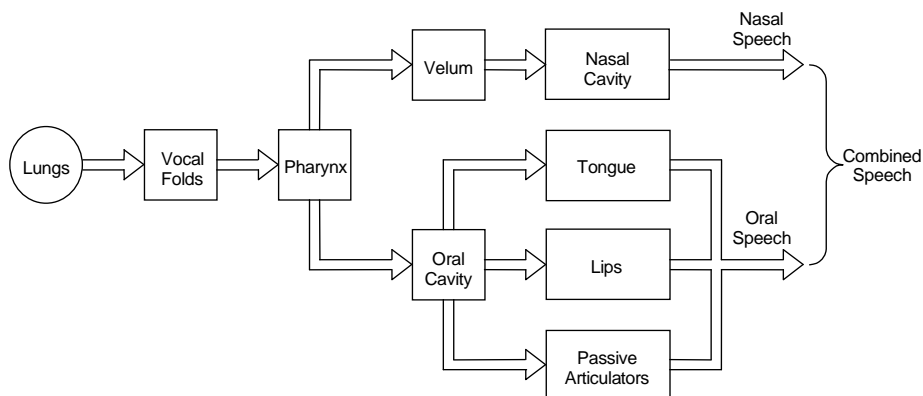


Figure 1: Graphical illustration of the speech production process.

The airstream to power the speech process is generated in the lungs and for many of the speech sounds the vocal folds are used to interrupt this flow of air, causing periodic pulses, or phonation. Differing length and mass of vocal folds lead to different fundamental frequencies of vibration: around 125 Hz in men, 200 Hz in women, and 300 Hz in children. During speech, the frequency of vibration alters, by varying the tension of the vocal folds, as pitch is changed in intonation.

The air pressure waves then pass through the pharynx. Its role in speech is that of a resonating cavity, the dimensions of which can be altered, e.g. shortened or lengthened, by raising or lowering the larynx. Articulators such as the velum, lips and tongue may further modify the airstream.

During normal respiration, the pharynx is coupled to the nasal cavity, as the velum is relaxed – this is also the case during the production of nasal consonants, e.g. /m/ where the pharyngeal, oral and nasal resonating cavities are involved. However, for the vast majority of the consonants of English, the nasal cavity is closed, by elevating the velum. The degree of closure necessary is dependent on the sound, and its phonetic context. The lips have three functions. First, they are a place of closure, as in /p/. Secondly, they can alter the size and shape of the oral resonance cavity, by altering lip shape, e.g. /Y/. Finally, they can act as a sound source, e.g. during /f/ where air passes through the gap between the upper incisors and the lower lip, under pressure, causing friction.

The most versatile of the articulators is the tongue, which is involved in the production of all vowels and the vast majority of consonants. Different sounds require different tongue configurations. By altering tongue position and shape, the size of the oral cavity and therefore its resonating characteristics, are changed. For example, with the voiceless alveolar plosive /t/ the *superior longitudinal* muscle enables the tongue to form a seal around the alveolar ridge and edges of the hard palate. The velum rises as the *levator palatini* contracts, and closes against the pharyngeal wall. Expiratory air builds up pressure in the oral cavity and this is released as the tongue rapidly comes away from the alveolar ridge.

Finally, the passive articulators, such as the teeth, hard palate and mandible, contribute to the articulation of speech sounds. For example, as mentioned above, the teeth are used with the lips in the production of some sounds. Additionally, the mandible acts to vary the size of the oral cavity, thus modifying its resonant properties.

When it is considered that the average rate of speech is up to 4 syllables per second, each of which can contain anything up to seven consonants and a vowel sound, the complexity of articulator movement becomes apparent. It has been estimated that over 100 muscles are involved in the speech process² and their controlled co-ordination requires around 140,000 neuromuscular events every second³. If the timing and/or position of articulators are not properly controlled, abnormal speech may occur.

SPEECH DISORDERS

Disorders of speech can be either developmental or acquired and may be either physical or neurological. For example, acquired dysarthria may result from neurological disorders, such as Parkinson's disease, motor neurone disease and stroke³. Dysarthria is the most common of the acquired disorders of speech and language⁴. Physical developmental speech problems arise from conditions such as cleft palate⁵. One in 700 babies is born with a cleft lip and/or palate; around half of these will require speech and language therapy⁶. Speech intelligibility problems in both developmental and acquired disorders result from difficulties in co-ordinating the articulatory system. For example, in cleft palate speech, several characteristics have been well established. Where tongue placement is impaired, there may be lateralisation or palatisation of sounds. Another common form of articulatory difficulty involves problems with the position of the velum during speech, causing nasalised and weak consonants, nasal emission or turbulence accompanying consonants or nasal fricatives replacing consonants. As well as being one of the hallmark characteristics of cleft palate speech, velopharyngeal incompetence (VPI) is common in those that have acquired dysarthria

ASSESSMENT TECHNIQUES

Assessment of speech defects is initially subjective, relying on the clinical judgement of the Speech & Language Therapist. This will involve both assessment of the intelligibility and quality of the patient's speech, and observation of the visible aspects of articulation (e.g. lip and some tongue movement). However, the majority of the articulators are not visible during speech. Additionally, there is a growing need for evidence-based intervention. Therefore, objective quantitative assessment is increasingly important. A number of individual instruments are available, which can achieve this.

- Videofluoroscopy – velum & tongue
- Nasendoscopy – velum and larynx
- Electrolaryngography/Laryngograph® – larynx & vocal folds
- Nasal anemometry/SNORS – velum
- Electropalatography/Linguagraph – tongue

VIDEOFLUOROSCOPY

Videofluoroscopy records moving x-ray images onto videotape. This provides a view of the velum and tongue during speech and provides useful information on the dynamics of these articulators. However, use of this technique is limited by the need to employ ionising radiation. Also, only a gross cross-sectional view is provided⁷.

NASENDOSCOPY

This technique utilises an endoscope, passed through the nares and nasal cavity, to image the velum, during speech⁸. The vocal folds can also be viewed in this way, and the use of a stroboscope allows the motion of the vocal folds to be observed. While providing very useful visual data, this technique is poorly tolerated by some patients⁹. Furthermore, both videofluoroscopy and nasendoscopy are very expensive techniques.

ELECTROLARYNGOGRAPHY

Laryngograph® is a commercial electroglottograph, developed and manufactured by Laryngograph Ltd, which measures vocal fold activity¹⁰. This is achieved by placing a pair of electrodes on the patient's neck, either side of the thyroid cartilage. By passing a small electric current through the vocal folds and measuring impedance changes, it is possible to detect their vibration as well as simple movements of the glottis.

NASAL ANEMOMETRY

Nasal anemometry measures nasal airflow during speech¹¹. This provides a measure of velopharyngeal closure. Whilst anemometry is an established technique in the UK, it has limitations such as sensitivity to speech loudness and slow response of the sensors. The University of Kent system, SNORS (Super Nasal Oral Ratiometry System), is also an anemometry system, which assesses velopharyngeal closure. However, unlike conventional nasal anemometers, SNORS measures both nasal & oral airflow, providing a much better assessment of the *degree* of closure^{9,12}. This is achieved by calculating the percentage of airflow that is nasal, thus virtually eliminating the volume factor. Also, SNORS uses fast sensors, which allow the very rapid movements of the velum to be detected. These sensors are bi-directional, thus allowing the system to differentiate between ingressive and egressive airflow.

ELECTROPALATOGRAPHY

Linguagraph is the University of Kent's clinical electropalatography system¹³. Electropalatography can determine tongue-palate contact, by using a special artificial palate, containing 62 electrodes on its tongue-facing surface¹⁴. A small electrical signal, fed to the patient, is conducted through the tongue to any touched electrodes and thence, via the electronics unit, to the computer where the tongue-palate

contact is displayed. The Palate used is that developed at the University of Reading, which has become the defacto standard in the UK. The palate is custom-made and simply clips to the upper teeth.

SYSTEM DESCRIPTION

SNORS+

SNORS+ encompasses all of the above instrumentation in a single PC based system. The software provides a user-friendly interface, which allows the simultaneous measurement of five key speech parameters. These parameters may be displayed as trend waveforms over time, or as 2D dynamic images. Synchronous audio and/or video recording and playback allows accurate identification of sound elements within the featured waveforms. Data from the various instruments are synchronously combined within the interface unit, which facilitates a single connection to the host computer's data acquisition card.

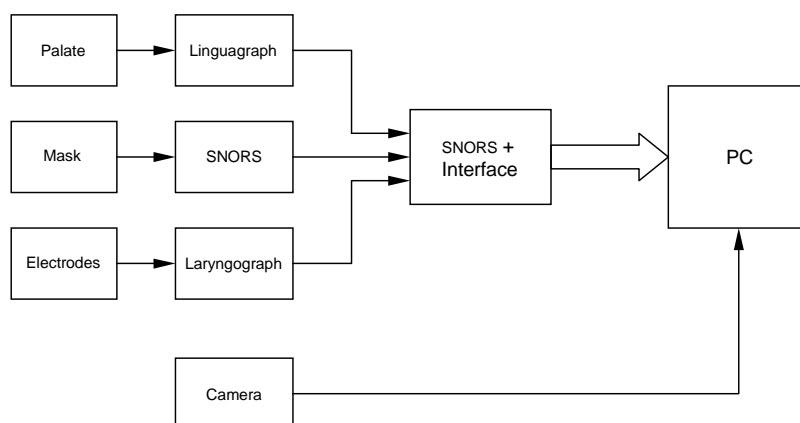


Figure 2: *A block diagram of SNORS+.*

THE USER INTERFACE

Written in C, SNORS+ has been specifically designed for a 32-bit Windows GUI interface, such as Windows® 95/98. Since its introduction, Microsoft Windows has emerged as the most popular graphical user interface environment¹⁵. Programs written for Windows have a consistent appearance and command structure and are often easier to learn and use than conventional DOS programs. In addition, Windows provides a wealth of built-in routines, that allow the use of menus, dialogue boxes, scroll bars, and other components of a friendly user interface. Windows programs run on a variety of hardware configurations allowing the programmer to treat peripherals such as the keyboard, mouse, video display and printer in a device independent manner. SNORS+ implements a multiple document interface (MDI), which is a specification that defines the standard user interface for applications written for Microsoft Windows. An MDI application enables the user to work with more than one document at a time. Each document is displayed in a separate child window within the client area of the application's main window. This allows the clinician to simultaneously view a large number of parameters in a variety of formats. The example illustrated in Figure 3, shows three child windows displaying a selection of speech parameters. The window to the right of the display features the following waveforms (top to bottom): speech intensity, nasal airflow, oral airflow and voicing intensity. A lateral Videofluoroscopic image of the velopharyngeal mechanism is illustrated in the top left-hand window. Finally, high-resolution speech and voicing waveforms are displayed in the bottom left-hand window.

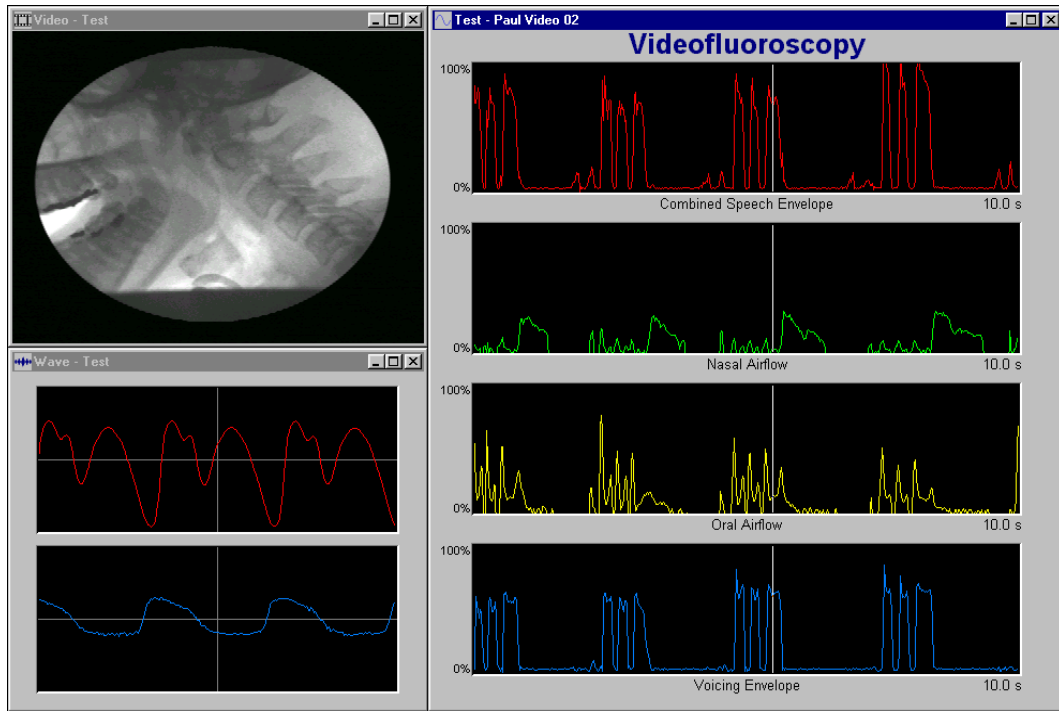


Figure 3: *The SNORS+ multiple document interface.*

The operation of SNORS+ can be partitioned into two main groups, real-time therapy and assessment/measurement of outcome.

REAL-TIME DISPLAY – THERAPY

In therapy mode, real-time displays are used to view the various speech parameters. Instantaneous 2D displays, showing single or multiple parameters, are generally the most useful for therapy, as they are uncluttered and easy to interpret. They provide clear visual feedback to the patient, allowing the effect of any speech corrections to be observed. All real-time displays are updated 100 times per second, which ensures that they are flicker free. The real-time Bar and electropalatography (EPG) displays are illustrated below.

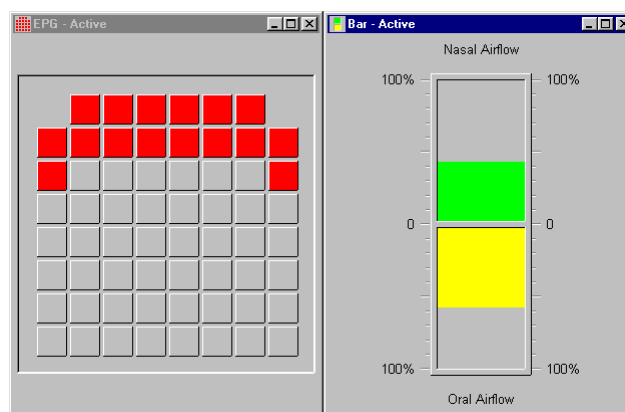


Figure 4: *The real-time Bar and EPG displays.*

The real-time Bar window is pictured in the right-hand display. The upper section of the window shows the amount of nasal airflow, as a green bar, which moves upwards away from the centre with increased nasal airflow. The lower section indicates oral airflow, as a yellow bar, which moves downwards away from the centre with increased oral airflow. Bar is a very useful biofeedback tool, as it provides a simple, clear display of nasal and oral airflow. This allows patients to monitor their airflow, make corrections, and observe the result. Bar is particularly useful when working with sustained sounds

or single phonemes. Other parameters can also be displayed, using Bar, e.g. voicing and sound intensity. A peak level indicator and target markers can be enabled, if required.

The left-hand display in Figure 4 shows a real-time EPG window. By asking the patient to produce certain sounds, the tongue-palate contact patterns can be observed on the display. Each of the small squares represents one of the 62 electrodes on the electropalatography palate, with those highlighted (coloured red on the computer display) representing areas of tongue-palate contact. In this case, contact is in the alveolar region, behind the incisors (top). To encourage the patient to produce certain tongue positions, model contact patterns may be retrieved from disk file. Figure 5 illustrates a subject's articulation of /s/ and the file model provided. The active EPG (left) and model (right) appear tiled on the screen.

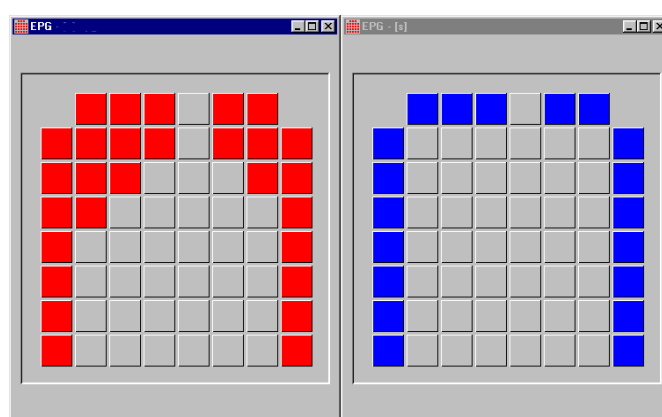


Figure 5: The real-time EPG and model EPG displays.

In addition to the instantaneous displays, real-time trend displays are also available. These show information about the dynamics of articulation. Traces scan across the display from left to right in a fashion similar to that of an oscilloscope. Up to nine traces may be viewed simultaneously. This display, termed “Scope”, is a useful therapy tool because it allows the patient to observe the dynamics of their speech during the utterance of complete words or phrases. The Scope traces may be frozen at any point during the scan, allowing the clinician to discuss the display with the patient. A three-channel Scope display is illustrated in Figure 6.

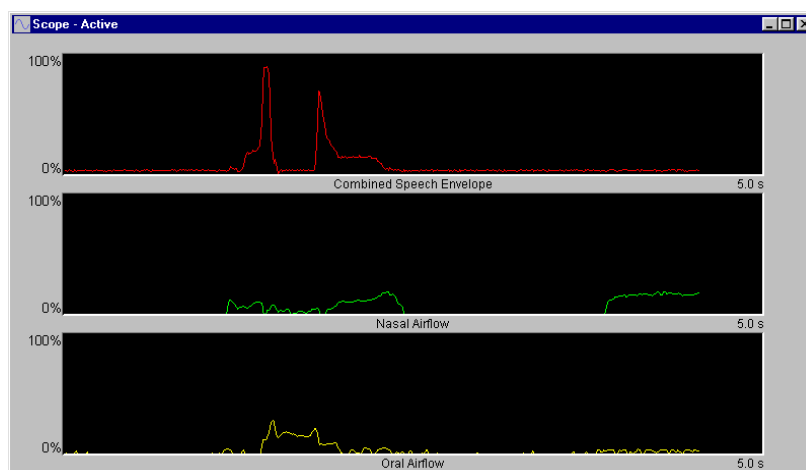


Figure 6: A three channel real-time Scope display

ANALYSIS DISPLAY – ASSESSMENT AND MEASUREMENT OF OUTCOME

To conduct a formal assessment, it is necessary to define a protocol that allows comparable measurements to be performed. The “Test” protocol defined in SNORS+ examines velopharyngeal closure only, and requires the patient to utter a number of words into the facemask, as prompted by the computer screen. The standard word list routinely used by the authors consists of the words: *begin, type,*

fight, seat, cheese, shoot, smoke, king, missing, end. The list provides a good selection of non-nasal and nasalised words ¹¹. The words appear individually on the computer screen in a large font. The time interval between each word is 2 seconds by default, but may be adjusted anywhere between one and ten seconds if required. The word lists are generated in a simple text editor launched within the package. Additional parameters, e.g. voicing and tongue-palate contact, may be selected and recorded. The particular combination of parameters and word list can be customised to suit a particular patient. Depending on the selected parameters, at the end of each test, a multiple window display is generated (see Figure 7).

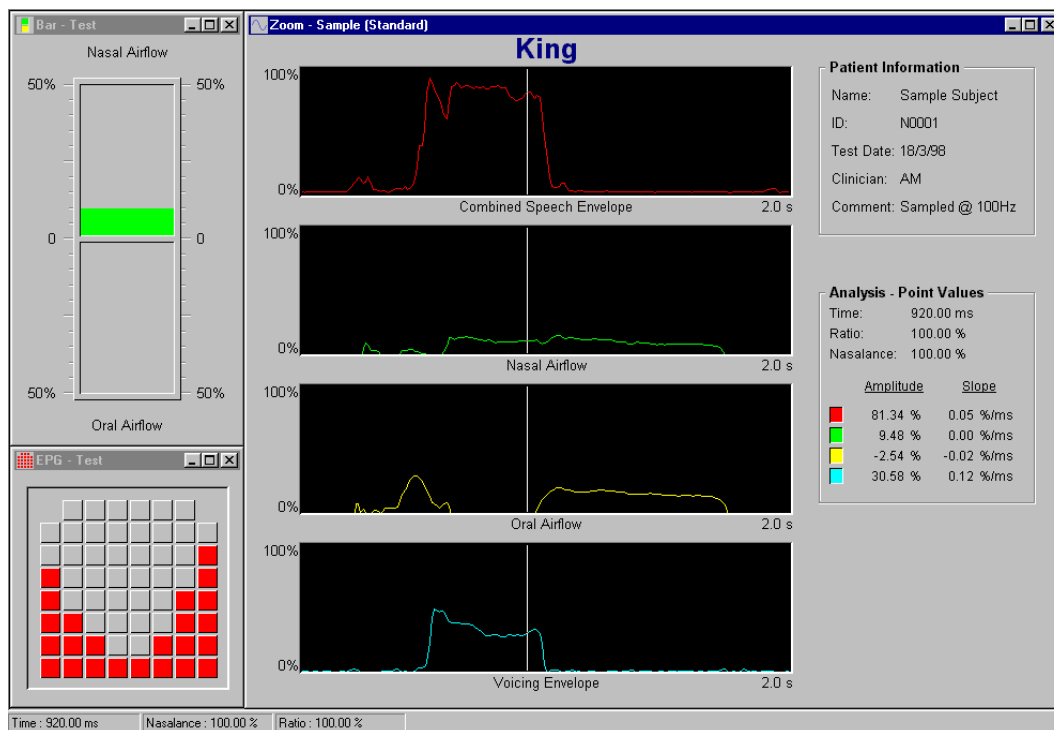


Figure 7: A zoomed Test display featuring EPG and Bar windows.

Analysis is performed on the main “Test Scope” window by positioning a cursor over areas of interest. A variety of waveform analyses, relating to cursor position, is displayed to the right of the window and also below on the status bar. As the cursor is positioned over the waveforms, any active 2D displays change to reflect the data at the current cursor position. In the example illustrated in Figure 7, tongue-palate contact and nasal/oral airflow is shown for a cursor positioned 0.92 seconds into the recording for the word King. Audio playback of the displayed speech waveforms may also be activated; this often makes the interpretation of the displayed waveforms easier. When a cursor is active, audio playback will commence from the cursor position. This is useful for isolating specific speech sounds. The audio data may also be viewed as high-resolution waveforms within a child window (see Figure 3, above).

SYSTEM INTEGRATION

The main advantage of SNORS+ over standalone instrumentation is its ability to measure and display the various speech parameters **simultaneously**, thus presenting the clinician with a comprehensive view of the speech process. It is therefore essential to the successful operation of SNORS+ that the signals obtained are properly synchronised, since loss of synchronisation may result in misinterpretation of the data.

DATA ACQUISITION

SNORS+ utilises three PC based data acquisition cards to capture the various parameters:

- Keithley Metrabyte® DAS-1202: SNORS, Laryngograph® and Linguagraph
- Creative Labs® Sound Blaster 16: SNORS audio and Laryngograph®
- Matrox Meteor® Frame Grabber: Videofluoroscopy or Nasendoscopy

The Keithley Metrabyte® DAS-1202

The DAS-1202 is a multifunction analogue and digital interface board that installs directly into an ISA expansion slot on the host computer. The DAS-1202 offers 16 single ended analogue inputs with 12-bit resolution at up to 100 ksamples/s. A three channel programmable timer provides timing for the analogue to digital conversions. Data transfer may take place in one of three ways: program control, interrupt service routine (ISR) or direct memory access (DMA). DMA mode has been selected for this application because it offers maximum transfer rates. Additionally, DMA operations execute as background tasks, which allow the application (and operating system) to execute other instructions while the DMA operation is in progress.

Analogue channels 0 through to 3 are used to capture SNORS data, i.e. nasal/oral airflow and nasal/oral speech envelopes. The analogue output from each airflow sensor passes through a 35 Hz low-pass filter prior to data acquisition. The purpose of this filter is to remove the unwanted voicing component of the airflow, while retaining the respiratory information relating to the velopharyngeal mechanism. The speech signals acquired by the DAS-1202 are primarily used as a reference for the airflow signals. The authors felt it beneficial to pass the envelope of the speech to the PC to convey this information, in order to provide a less cluttered display. The use of enveloped signals also allows for a lower sampling frequency, eliminating the problems of both sampling and screen aliasing. The envelope generator was implemented using an active half-wave rectifier followed by a 35 Hz low-pass filter.

The signal obtained from the Laryngograph® unit is processed in a similar manner to that of the speech signals, i.e. half wave rectified and low-pass filtered. Again this produces a less cluttered display. Laryngograph® data is sampled through the fourth analogue input channel of the DAS-1202.

Due to the relatively low frequency of the signals described above, acquisition rates as low as 100 Hz may be used to sample the data without introducing significant distortion. However, to ensure flexibility, the system supports sample rates of up to 10 kHz per channel. The DAS-1202 programmable timer determines the sample period for each data acquisition sequence.

The remaining eleven analogue inputs (5 through to 15) are used to acquire Linguagraph data (see Figure 8). The signals obtained from the Linguagraph system are derived from the 62 electrodes embedded within the artificial palate. The electrodes act as switch contacts when touched by the tongue, which conducts an 800 Hz body clock signal. The body clock, generated by the DAS-1202 programmable timer, is opto-isolated and applied to the patient via an electrode worn on the wrist. Using the programmable timer to generate the body clock ensures the Linguagraph data is synchronised with the SNORS and Laryngograph® data. Due to the large number of signal lines obtained from the artificial palate, it is impractical to sample each one simultaneously. Therefore, to limit the required number of sample lines, eight, 8:1 analogue multiplexers have been employed. These present only eight signals to the data acquisition card at any one time. The multiplexer select lines are generated by a 3-Bit step down counter attached to the body clock. The resulting signal, sampled by the DAS-1202, is a digital representation of eight of the contacts on the artificial palate. Thus, a total of eight scans are required to capture one frame of Linguagraph data. Each 8-Bit scan is uniquely identified by a 3-Bit address derived from the multiplexer select lines. Using this address, the software can accurately reconstruct each frame.

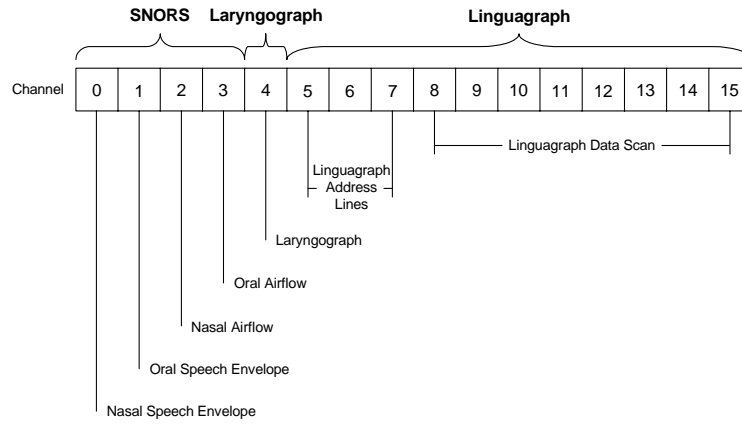


Figure 8: *A typical DAS-1202 data scan*

It is expected that future Linguagraph designs will measure tongue/electrode proximity. This method should increase the overall sensitivity of the system as the tongue often approaches the palate electrodes but fails to make contact. To facilitate this approach, an analogue system would be required. Therefore, to ensure future compatibility, analogue sampling techniques have been employed here.

Creative Labs® Sound Blaster 16

SNORS+ uses the Creative Labs® 16-bit sound card to record and playback audio. The sound card is controlled via the Windows Media Control Interface (MCI) which provides standard commands for recording and playing multimedia resource files. The recording quality can be either CD (44 kHz), radio (22 kHz), or telephone (11 kHz), with a 16-bit resolution. The combined unfiltered speech signals (nasal plus oral) are recorded on the left channel of the sound card, with the right channel recording the unfiltered Laryngograph® signal. As with the DAS-1202 data acquisition card, the sampled data are transferred using DMA techniques. The acquired audio data are synchronised with the DAS-1200 data, to within a nominal 2ms.

The Matrox Meteor® Frame Grabber

Video images are acquired with the Matrox Meteor® frame grabber card. The Meteor, which installs directly into a PCI expansion slot on the host computer, is capable of recording colour or monochrome images at rates of up to 30 frames a second. At present, however, SNORS+ only captures monochromatic images at rates of up to 25 frames a second. Image data are transferred directly to the host computer's memory using interrupt service routines. The card supports the following camera formats: PAL, SECAM, CCIR, NTSC and RS170. The Matrox Imaging Library (MIL) package is used to program the meteor, which enables SNORS+ to grab and display synchronised images, enhance images and zoom. Video synchronisation with the DAS-1200 data is to the nearest frame (i.e. 40ms, at 25 fps).

THE TEST DATA ACQUISITION THREAD

During the Test sequence, a single data acquisition thread controls the execution of all data capture routines. A thread is a section of code that can run concurrently with other threads (or processes) within the system. Threads allow the CPU to appear to be executing two sections of code simultaneously. When created, each thread is given a priority, which governs the amount of CPU time the thread is allocated. Placing the data acquisition routines in a high priority thread ensures they take precedence over many other instructions awaiting execution. However, due to the multitasking nature of threads, during periods of inactivity the data acquisition thread relinquishes the CPU, thus allowing the execution of other instructions, i.e. operating system tasks. Threads ensure that the Windows® operating system remains responsive, even during the busy data acquisition period.

The structure of the test data acquisition thread is illustrated below in pseudo code.

```
BEGIN
    if enabled initiate audio recording;
    initiate DAS-1202 data acquisition;
    if enabled record initial video frame;
    display first word in list;

    REPEAT
        check number of DAS-1202 samples;
        if next video frame due initiate next frame recording;
        if word data captured prompt user with next word;
    UNTIL
        all words in list have been displayed;

    if enabled stop audio recording;
    stop DAS-1202 data acquisition;
END
```

Once captured, the synchronised data are transferred to the appropriate functions for processing and eventual display. Similar data capture techniques have been employed for the real-time therapy modules.

RESULTS & DISCUSSION

We present sample results to illustrate the use of SNORS+. As this is a clinical instrument, results from two patients are compared with those from a normal subject. This graphic demonstration of the differences and the interaction of the articulators allows a clear evaluation of the patients' difficulties and their cause. For clarity, the results are presented as simple printouts of the SNORS+ information.

PATIENT BB

BB is a 42 year old female, with lupus erythematosus plus Sjorgren syndrome and vesiculation of the tongue. As a result of limited function of the palatopharyngeal muscles, BB has a moderate oral / severe pharyngeal dysphagia and hypernasal speech. She expressed concern about the possibility of vocal fold damage, as a result of a temporary tracheostomy. Hence, she was assessed using SNORS and Laryngograph. Thus, the results show four traces: the sound envelope, nasal airflow, oral airflow and voicing envelope.

Figure 9 shows the results from a normal subject, saying the word "cheese". Looking at the speech envelope, we initially see a low-level sound, representing the affricate, /tʃ/ ("ch"). The sound then peaks, at the start of the vowel, /i/ ("ee"), and continues at a fairly high level during the vowel and the final consonant, /z/ ("se").

Looking at the airflow, there is virtually zero nasal airflow throughout the word, except for a small peak as the affricate is released (due to flexing of the velum). The oral airflow, however, peaks considerably during the release of the affricate, /tʃ/ ("ch"), and remains at a fairly high level during the vowel and the final consonant. After the word, both oral and nasal airflows rise, due to exhalation.

The voicing starts with the vowel and remains at an almost constant level during the vowel and the final consonant. This is as expected as the affricate is unvoiced, but the vowel and final consonant are voiced.

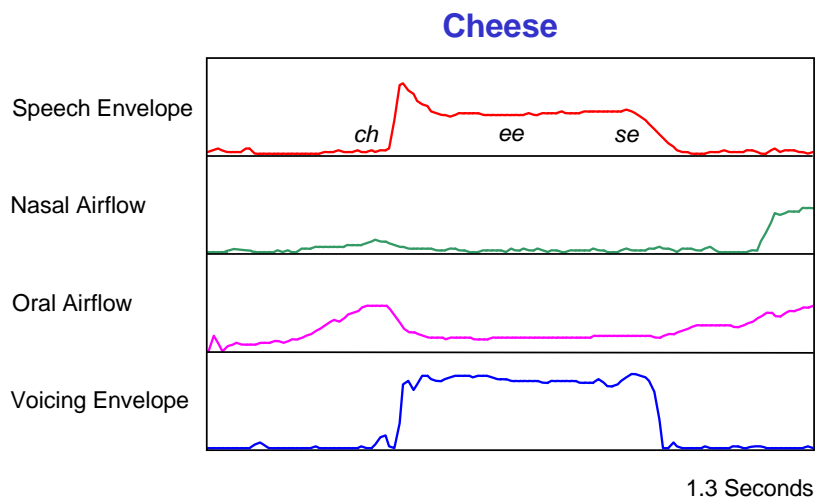


Figure 9: Normal waveforms for the word “cheese”

Considering the results for patient BB (Figure 10), we see that the resulting speech envelope is reasonably normal. If we look at the airflow data, however, we see that the nasal airflow exceeds the oral values throughout the word, confirming the escape of nasal airflow and indicating that the velopharyngeal port is failing to close. This results in BB’s hypernasal speech. Again, there is both oral and nasal airflow after the word as the subject exhales.

The voicing envelope is fairly normal, persisting during the vowel and the final consonant, showing that voicing is well co-ordinated. There is some amplitude variation, but this is probably due to pressure fluctuations, resulting from the velopharyngeal incompetence. The high-resolution voicing waveform (not shown) was also normal. Hence, this assessment showed no evidence of vocal fold damage, though further ENT investigation would be necessary to confirm this.

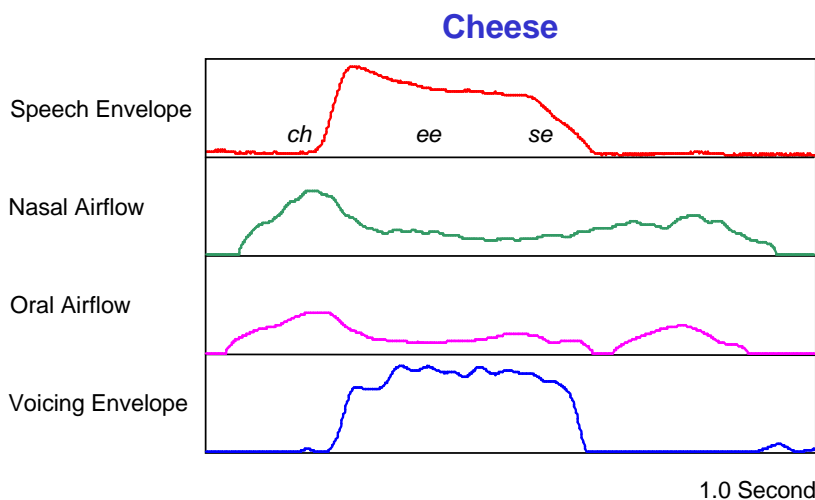


Figure 10: Waveforms for the word “cheese” for patient BB

PATIENT CP

CP is an 81-year-old male, who had a CVA in 1994, resulting in a moderate spastic dysarthria. Additionally, he had basal cell carcinoma, which caused a fistula in the region of the nasal bone, close to the left eye socket. He received radiotherapy in the treatment of further lesions around his mouth. His speech difficulties were treated with conventional Speech and Language Therapy and electropalatography therapy.

Spastic dysarthria can affect any of the articulators and co-ordination of these. CP had responded well to electropalatography therapy, but his speech difficulties remained. He was, therefore, assessed using SNORS, Laryngograph and Linguagraph. This allows us to view the action of the velum, larynx and tongue. Hence, the results show the four traces as before, plus snapshots of tongue-palate contact at two points of interest.

Figure 11 shows the results for a normal subject, for the word “seat”. Looking at the sound envelope, there is initially a very low-level sound for the fricative, /s/. This is followed by a short, high-level period, during the vowel, /ɪ/ (“ea”). There is then a silence, during closure, followed by a small peak for the release of the final /t/.

The nasal airflow is very small, except for the usual exhalation at the end of the word. The oral airflow is high during the /s/, but falls to a low level during the vowel. This is because the resonant vowel sound results in little DC airflow, comprising mainly the AC perturbations producing the sound pressure wave. The oral airflow finally peaks as the plosive /t/ is released.

Voicing is apparent only during the vowel. This is as expected, as the fricative /s/ and the plosive /t/ are unvoiced.

The two snapshots show the tongue-palate contact patterns (highlighted squares) during the production of the fricative /s/ (snapshot A) and during the closure for the plosive /t/ (snapshot B). For the /s/ we see that the tongue contacts the hard palate around the dental arch and alveolar ridge, leaving just a small groove, through which pressurised air is blown, causing friction, to produce the sound. The pattern for the /t/ is similar, except there is no groove. Hence, there is a complete seal, which allows air pressure to build up. It is the transient airflow resulting from the release of this pressure, which causes the sound.

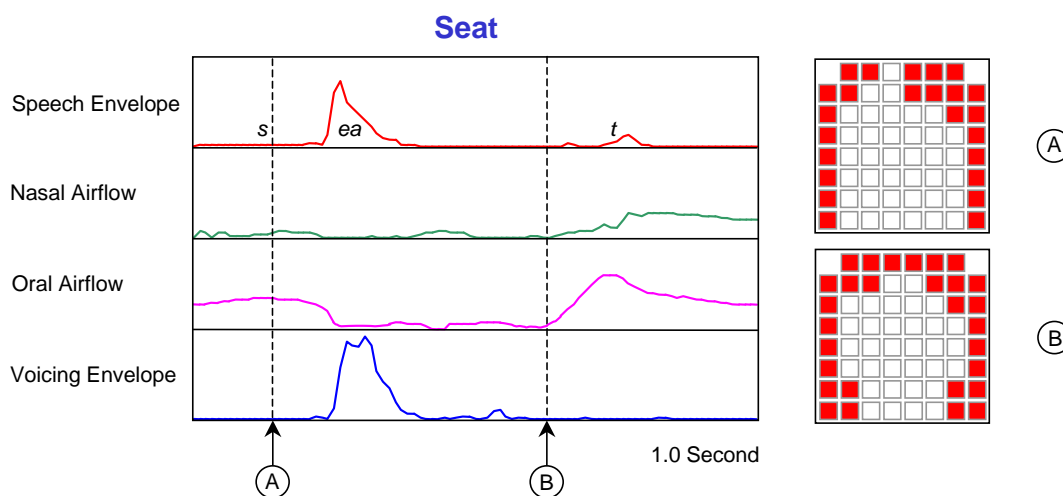


Figure 11: Normal waveforms for the word “seat”

For patient CP (Figure 12), the sound, airflow and voicing waveforms are all similar to the normal traces. This suggests that CP has adequate control of his velum and vocal folds. The tongue-palate snapshots, however, are abnormal. Considering the /s/ (snapshot A), while there is good contact down both sides, CP fails to make contact in the front, alveolar region (top). This shows that he lacks the fine tongue control necessary to create the groove and articulate this sound. For the closure of the /t/ (snapshot B), CP makes effective closure, but the contact is very heavy, compared to the normal example. Again, this illustrates his lack of fine control.

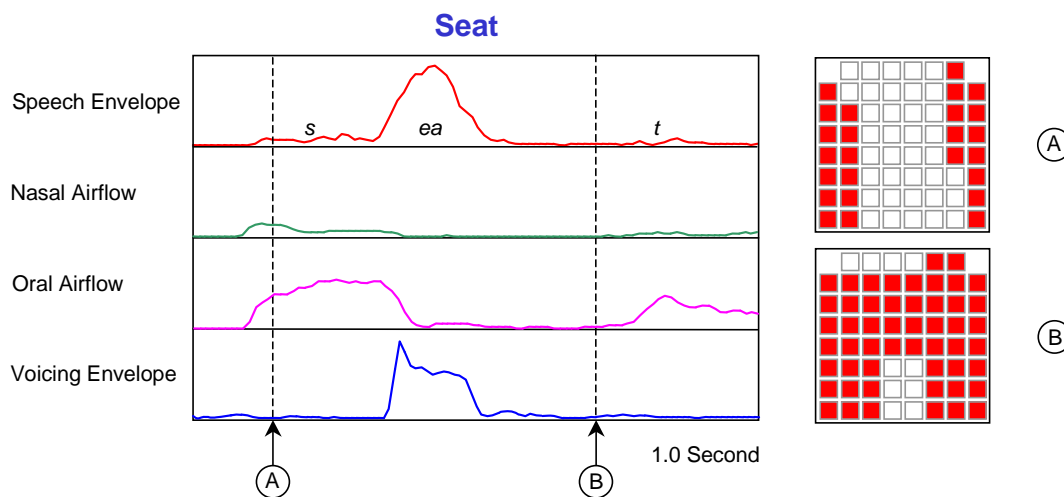


Figure 12: Waveforms for the word “seat” for patient CP

CONCLUSIONS

SNORS+ is a clinical, user-friendly, instrument, which has been developed for measurement of the articulators, during speech. This affordable system allows simultaneous, multiparameter assessment or real-time display. It thus enables objective analysis of speech production and biofeedback from the articulators.

Preliminary results suggest that such an instrument will prove extremely useful in the assessment and management of many speech disorders. SNORS+ allows objective assessment of the function and coordination of key articulators. Appropriate targeting of therapy is therefore possible. It provides a visual feedback, which assists in therapy and is very motivating for therapist and patient. And it gives an objective measurement of outcome, which is an increasingly important consideration.

The system has been used with a small number of patients, with a range of speech difficulties. Further, more extensive trials are planned and it is hoped that this will enable normative data to be obtained and a clinical evaluation of SNORS+. Thus, the full diagnostic and therapeutic potential of the equipment will be established.

ACKNOWLEDGEMENTS

The authors would like to thank Laryngograph Ltd, whose collaboration is invaluable. We would also like to thank the patients who have acted as willing guinea pigs in the preliminary evaluation of SNORS+.

REFERENCES

- 1) Borden, G.J., Harris, K.S., and Rafael L.J., *Speech Science Primer*, 3rd edn. Williams and Wilkins, Baltimore, 1994.
- 2) Lenneberg, E.H., *Biological Foundations of Language*. John Wiley and Sons, New York, 1967.
- 3) Darley, F.L., Aronson, A.E., and Brown, J.R., *Motor Speech Disorders*. WB Saunders, Philadelphia, 1975.
- 4) Enderby, P.M., and Emerson, J., Speech and Language Therapy: Does it work? *British Medical Journal*, 1996, **312**, 1655-1658.

- 5) Harding, A. and Grunwell, P., Characteristics of cleft palate speech. *European Journal of Disorders of Communication*, 1996, **31**, 331-357.
- 6) Sandy, J., Williams, A., Mildenhall, S., Murphy, T., Bearn, D., Shaw, B., Sell, D., Delvin, B. and Murray, J., The clinical standards advisory group (CSAG) cleft lip and palate study. *British Journal of Orthodontics*, 1998, **25**(1), 21-30.
- 7) Kuehn, D., Assessment of resonance disorders. In: Lass, N., McReynolds, L., Northern, J., Yoder, D. (Eds) *Speech Language and Hearing: Vol 11 Pathologies of Speech and Language*. WB Saunders, Philadelphia, 1982, 499-524.
- 8) Pigott, R., Benson, J., White, F., Nasendoscopy in the diagnosis of velopharyngeal incompetence. *Plastic and Reconstructive Surgery*, 1969, **43**, 141-147.
- 9) McLean, C.C., Kelly, S.W., Manley, M.C.G., An instrument for the non-invasive objective assessment of velar function during speech, *Medical Engineering & Physics*, 1997, **19**(1), 7-14.
- 10) Abberton, E. and Fourcin, A., Electrolaryngography. In: Ball, M.J. & Code, C. (Eds), *Instrumental Clinical Phonetics*, Whurr, London, 1997, 119-148.
- 11) Ellis, R.E., Flack, F.C., Curle, H.J. and Selly, W.G., A system for the assessment of nasal airflow during speech. *British Journal of Disorders of Communication*, 1978, **13**(1), 31-40.
- 12) Main, A., Kelly, S. and Manley, G., Instrumental assessment and treatment of hypernasality, following maxillofacial surgery, using SNORS: a single case study. *International Journal of Language & Communication Disorders*, 1999, **34**(2), 223-238.
- 13) Main, A., Kelly, S.W. and Manley, M.C.G., Teaching the tongue and looking at listening – the use of electropalatography with a patient with severe ataxic dysarthria. *Royal College of Speech and Language Therapists Bulletin*, 1997, **November**, 8-9.
- 14) Hardcastle, W., Jones, W., Knight, C., Trudgeon, A. and Calder, G., New developments in electropalatography: a state-of-the-art report. *Clinical Linguistics and Phonetics*, 1989, **3**, 1-38.
- 15) Petzold, C. and Yao, P., *Programming Windows 95*. Microsoft Press, Washington, 1996.