Speech rehabilitation assessment using the SNORS+ system incorporating EMG: phonetic analysis

MSc. Rosario Dominguez-Camara and Eur. Ing. Stephen W. Kelly
Department of Electronics
University of Kent at Canterbury
CT2 7NT, Canterbury, Kent. United Kingdom
Telephone: +44(01227) 82-7621, Fax: +44(01227) 45-6084. E-mail: rd9@ukc.ac.uk

Abstract

This paper describes the utilisation of the SNORS+ system for speech therapy as well as the addition of a new parameter for speech disorder understanding. Multiparametric assessment is one of the main aids to speech rehabilitation because it allows objective evaluation of the function and co-ordination of key articulators involved during speech production.

Keywords: SNORS+, Electromyographic Signal (EMG), Multiparameter, Speech Disorders.

1. Introduction

Speech is the result of a highly complex and versatile system of co-ordinated muscular movements [1]. The main organs of the human body responsible for producing speech are the lungs, larynx, pharynx, nose, mouth, lips, teeth, jaw, vocal cords and velum. Muscular force, to expel air from the lungs, provides the source of energy. The airflow is then modulated in various ways to produce components of acoustic power in the audio frequency range. The properties of the resultant sound are modified by the rest of the vocal organs to produce speech. For each sound, there is a particular position for each of the vocal tract articulators: vocal cords, tongue, etc. Fig. 1 shows the speech production block diagram.

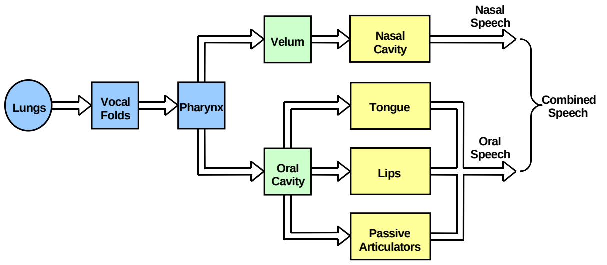

Figure 1. Speech production process block diagram.

Speech sounds comprise periodic and/or aperiodic sound waves. Periodic sound waves are produced by the vocal folds, which lie in the larynx, and then modified, by other articulators. On the other hand, allowing air to pass through the open vocal folds into the vocal tract to be modified by the articulators, creating friction produces aperiodic sound waves. Additional complex sounds are produced by the combination of the two methods.

The use of instrumentation has increased enormously within the medical field over the past few decades. SNORS+ is a multiparameter workstation developed to assess speech disorders, Fig. 2.

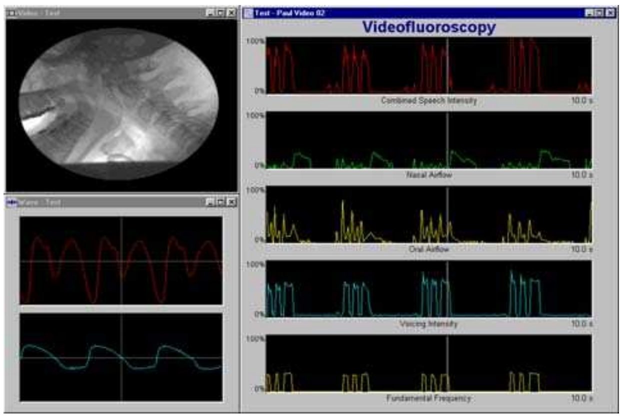

Figure 2. A SNORS+ parameters screen display. The window to the right of the display features the following waveforms (top to bottom): speech intensity, nasal and oral airflow, voicing intensity and fundamental frequency. A lateral Videofluoroscopic image of the velopharyngeal mechanism is illustrated in the top left-hand window. Finally, high-resolution speech and voicing waveforms are displayed in the bottom left-hand window.

The SNORS+ unit encompasses several speech evaluation techniques in a single program. Hardware and software provide a friendly interface between both clinician and patient allowing the simultaneous measurement of five key speech parameters: respiration, larynx excitation, velopharyngeal closure, tongue-palate contact and speech outcome, Fig. 3. In addition, audio playback provides accurate identification of the recorded multiparameter data, and a synchronised video input enables simultaneous use with well-established imaging techniques [2].

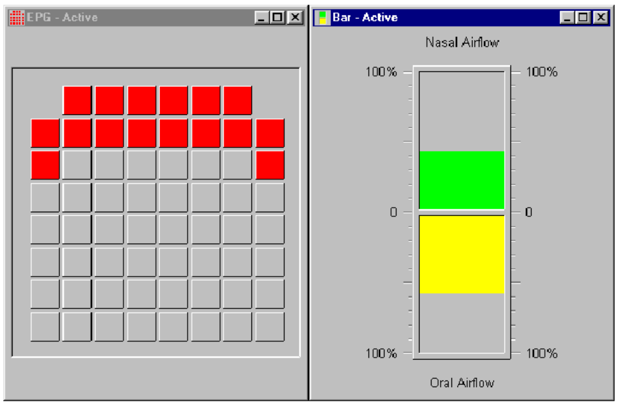

Figure 3. EPG and Bar child window results after an assessment. The EPG corresponds to the /l/ pronunciation and bar results show a nasal-oral word pronunciation.

The plan to introduce further important parameters to this therapeutic unit is described in this article. The current system assesses movements that result from muscular contraction. For example, movements of the tongue are assessed by measuring the contact points between the tongue and the hard palate. When these movements are incorrect or absent, it is not possible to determine whether this is due to weak or mistimed muscular effort, or structural or neurological problems. By simultaneously recording and analysing electromyographic signals, this project aims to establish the relationship between muscular contraction and resulting movements, during speech. An additional benefit will be that movements that are not currently measured, such as lip protrusion, will now be able to be assessed.

Disorders of speech can be either developmental or acquired and may be either physical or neurological. Some of the most common speech disorders are dysarthria, apraxia and stuttering [3]. Physical developmental speech problems arise from conditions such as cleft palate or lip [4].

Speech intelligibility problems, in both developmental and acquired disorders, are often the result of non co-ordination in the articulatory system.

For instance, dysarthria is a speech disorder that is due to a weakness or non-coordination of the speech muscles. As a result speech is slow, weak, imprecise or uncoordinated.

Apraxia is a motor disorder, in which voluntary movement is impaired without muscle weakness. The ability to select and sequence movements is impaired. Oral apraxia affects the ability to move the muscles of the mouth for non-speech purposes. Someone with oral apraxia would have trouble coughing, swallowing, wiggling his or her tongue or blowing a kiss when asked to do so.

Normal speech is defined as a fluent, smooth, forward-moving, unhesitant and effortless speech. A "dysfluency" is any break in fluent speech. Everyone has dysfluencies from time to time but stuttering is speech that has more non-fluency than is considered average. Dysfluencies that are more characteristic of stuttering include sound or syllable repetition, prolongation (unnatural stretching out of sound), and blocks (sound gets stuck and can not come out).

Cleft palate and lip is a common congenital abnormality occurring in 1 in 600 births and can be presented in a wide variety of forms and combinations. Cleft lip ranges from notching of the lip to a complete cleft, involving the floor of the nose and may be associated with a cleft of the primary palate (alveolus/pre-maxilla) and with clefts of the secondary palates (hard and soft palate), Fig. 4.

Figure 4. Congenital cleft lip, 3 months baby born.

Clefts may be unilateral or bilateral, complete, incomplete, or microform. Cleft palate may occur in isolation, may be unilateral or bilateral, and ranges from a bifid uvula to a complete cleft of hard and soft palates.

A person with cleft lip and/or palate can experience diverse types of speech and language problems. A cleft palate involves an opening or split in the oral structures that can be surgically repaired at a very early age. Although the surgeon may be successful in closing the cleft, it is often possible that difficulties with speech may be developed.

Problems with language delay can occur due to a child's inability to produce certain sounds. Compensatory speech patterns can develop and be very difficult for adults to understand. This in turn affects the person's interaction with adults and children.

Articulation can also be affected by the cleft palate if the soft palate cannot make contact with the pharyngeal walls (wall at the back of the throat). Airflow will leak through the nose during speech (hypernasal speech). This can affect consonants sound "p", "b", "t", "d", "s", "sh", "ch", and "f" if the seal cannot be maintained. As the air escapes, the sounds will become distorted.

The voice quality can indicate signs of hypernasality and other problems with resonance. In addition to problems with language and articulation, repeated ear infections and dental issues can further complicate speech and language of a cleft palate patient. Often, the early and late eruption of teeth, in conjunction with the collapse of the upper jaw further complicate the cleft lip/palate child's ability to develop speech and language skills in the traditional method.

The aim of cleft surgery is the restoration of normal anatomy and the promotion of normal growth and development of all structures affected by the cleft. Emphasis must be placed on the restoration of muscle continuity whether of the lip and/or nose, or the soft palate

After lip or palate surgery takes effect, it is desirable to assess the new movements of mouth, lips and velum, as well as the progress of the tongue's movements. It is here where the new system, incorporating muscle activity analysis, proves its usefulness. These movements can be assessed and analysed by the study of the electromyographic signal produced by the muscles involved, such as those of the tongue, velum and lips.

When the body performs any activity, the muscles produce an electrical current, which is proportional to the level of that activity. Electromyography (EMG) is the study of that muscle activity, in the same way that Electrocardiography (ECG) is the study of cardiac activity.

There are two basic methods of EMG measurements – intramuscular EMG and surface EMG. The former involves placement of a needle or hooked-wire electrode, into the muscle to be measured, fig. 5. This kind of measurement is well established in the diagnosis of neuromuscular disorders. It provides the "cleanest" signal or recording from a single muscle.

The latter involves detection of electrical activity within a muscle by placing surface electrodes, fig. 6, onto the skin, directly above the muscle (or muscles) concerned. It is therefore non-invasive, and as such it is relatively quick and simple to take a measurement. Their use seems justified in clinical research and treatment, due to the relative ease of application, reduced discomfort to the patient, and non-invasive nature [3]. There are, in the same way, two kinds of electrodes, disposable or re-usable.

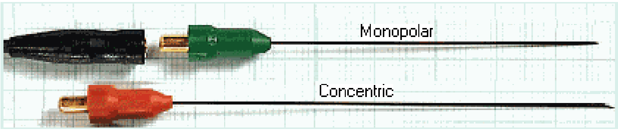

Figure 5. Needle Electrodes (Source: SLE Ltd).

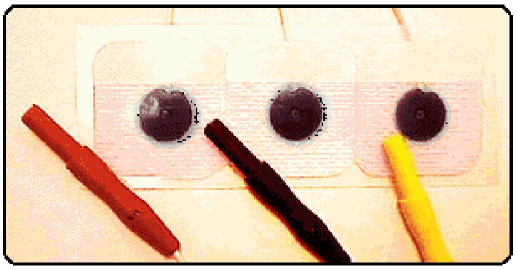

Figure 6. Surface electrodes (Source: SLE Ltd).

2. Anatomical Background

We aim to measure the electromyographic signal produced by the *Orbicularis Oris* muscle, and all those fibres derived from the other facial muscles, which are inserted into the lips. A considerable number of these fibres are derived from the *Buccinator* and form the deeper stratum of the *Orbicularis Oris*. Fig. 7 shows the structure and ramifications of the *Orbicularis Oris* muscle.

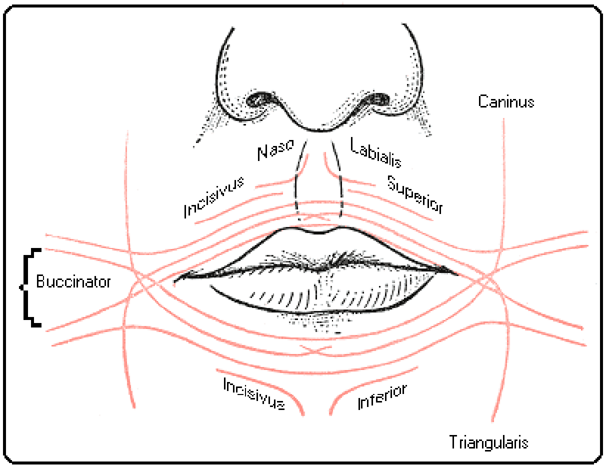

Figure 7. The *Orbicularis Oris* and its fibers.

The *Orbicularis Oris* originates in the near midline on the anterior surface of the *maxilla* and mandible and *modiolus* at the angle of the mouth. Its insertion is in the *Mucous membrane* of the margin of the lips and combines itself with the *buccinator* at *modiolus*. Its action is to narrow the orifice of the mouth, it purses the lips and puckers the lips' edges [5].

The upper lip is composed of the *Orbicularis Oris* muscle covered by skin on the outside and *Mucous membrane* on the inside. Many other muscles are attached to the *Orbicularis Oris* muscle, and together work to create the movements and to create the power required for speaking, eating and forming facial expressions. The *Orbicularis Oris* muscle constitutes the major bulk of the upper and lower lips.

The first aim is to measure those electromyographic signals produced by the lips during normal speech, and to verify whether is possible to obtain vowel and consonant patterns, closure and open patterns

The use of surface EMG biofeedback in speech was first described by Draizar in 1984 [6].

3. Procedure

Two electrodes are placed over the muscle to be assessed, over the lips, and a ground electrode is placed on a nearby site on the subject's body, on the *masseter* muscle [7]. The electrodes are connected to the amplifier via an electrode cable. The signal is amplified and then displayed on a Laptop computer screen, via a data acquisition card. Preparation of the electrode site and placement of the electrodes are a critical aspect when recording EMG. Poor preparation of the electrode site can result in high impedance, which will reduce the biological signal.

The electromyographic signal has inherent features. It is well established that the amplitude of the EMG signal is stochastic or random in nature and can be reasonably represented by a Gausian distribution function.

The amplitude of the signal can range from 0 to 10mV (peak to peak) or 0 to 1.5 mV (RMS) depending on the anatomical and physiological features of the muscle to be measured. The usable energy of the signal is limited to the 0 to 500 Hz frequency range, depending on the kind of contraction to be measured, with the dominant energy being in the 50-150 Hz range. Usable signals are those with energy above the electrical noise level [8].

The EMG signal associated with movement of the lips will be small and can be easily confused with noise; hence it is necessary to employ high quality filtering, amplification and conditioning circuits.

Having extracted the electromyographic signal, the procedure will be to identify specific features, related to individual lip forms occurring during speech. Finally, synchronisation of the EMG data with the other parameters such as nasal/oral airflow, EPG, videofluoroscopy and laryngography will be necessary.

4. Preliminary results

The initial measurements were those corresponding to lip protrusion and lip stretching. Specific measurements have been based on two words normally analysed by the SNORS+ system: SMOKE and CHEESE. Separate recordings have been made using both the SNORS+ system and our EMG analysis system.

First we will proceed to analyse the results given by the SNORS+ system for both words – "smoke" and "cheese". Our analysis will consist of an explanation of what is happening at every single phonetic sound.

The first word to be analysed is SMOKE, Fig. 8.

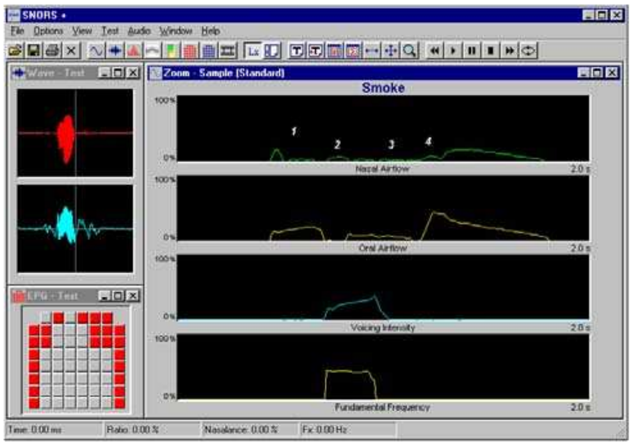

Figure 8. SMOKE phonetic analysis.

We are able to see that within this word there are four main stages. Stage (1): Corresponds to the phonetic sound /s/. We can observe there is oral airflow but not nasal; this is because the /s/ sound is an oral fricative sound [9]. This means that the airflow becomes turbulent in the region of constriction in the vocal tract. In this case the restriction is formed between the tongue and the hard palate; the tongue-palate contact pattern for this sound is seen in the EPG window at the bottom left of the display. There is no vocal cord vibration, which is why we do not get any signal for voicing intensity. Stage (2), during the /m/ sound, we find no oral airflow but nasal flow and vibration of the vocal cords. Next stage (3), as the /ŒX/ ("ow") sound starts appearing, the oral airflow increments for the diphthong, and the nasal airflow reduces to almost zero. Finally stage (4), one last pulse of oral airflow, corresponding to the /k/ sound, as this phonetic sound is an unvoiced stop consonant [9].

The second word CHEESE presents 3 stages, fig. 9.

Figure 9. CHEESE phonetic analysis.

It begins with the /tV/ ("ch") oral airflow sound (1), similar to the /s/ sound in the word smoke. It is an unvoiced africative sound, in which the constriction is located at the front of the vocal tract [9]. During stage (2), /1/ ("ee") sound production, there is oral airflow but not nasal and the vocal cord activity can be seen in the voicing intensity window. The last stage (3), comprises a /s/ sound which is a purely oral sound. The voicing intensity shows some signal due to the remaining vibration from the /1/ sound, but this dies away during the sound.

Now we will proceed with the EMG analysis for the same two words. At first the word "Smoke", Fig. 10.

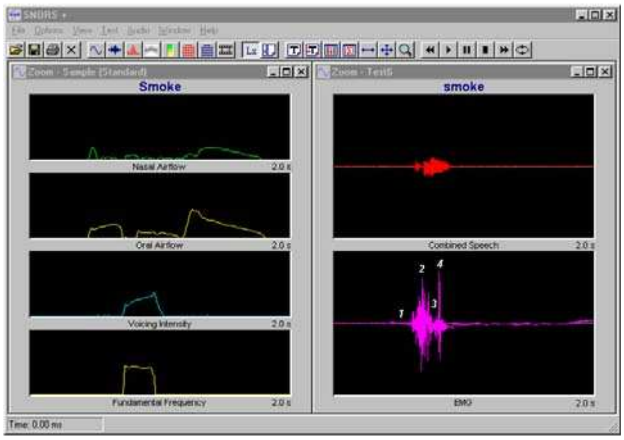

Figure 10. SMOKE EMG analysis.

With the EMG analysis it is possible to see the same four stages. The most important information is that during /s/ sound, stage (1), there is no EMG signal as there are no muscular movements (there is no lip protrusion). During the /m/, and /k/ sounds, stages (2) and (4), we can see there is a high level of muscular activity, due to the lips' protrusion to form the sound correctly. During the /ŒX/ sound the muscular signal is kept at a medium level as there is some lip movement needed to form this sound.

The second word "Cheese" posses the same three stages as previously, Fig. 11.

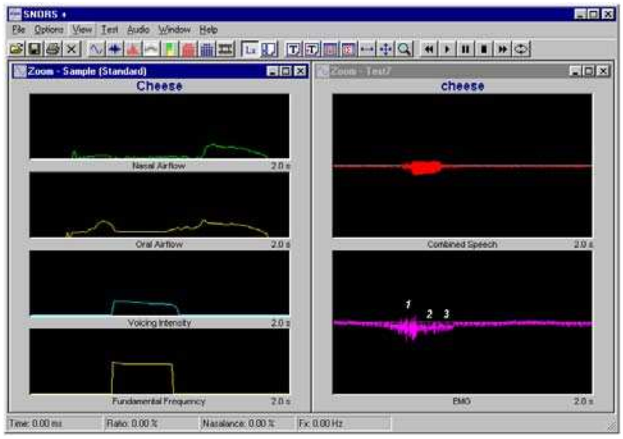

Figure 11. CHEESE EMG analysis.

During the whole word pronunciation there is no significant muscular presence. This is because the word cheese corresponds to a word in which the *Orbicularis Oris* is almost totally relaxed. As a result, we obtain only small amplitudes in the EMG signal.

5. Conclusions

The SNORS+ system is a clinical, user-friendly, instrument, which has been developed for measurement of the articulators, during speech production.

The workstation allows simultaneous, multiparameter assessment, enables objective analysis of speech production provides and biofeedback from the articulators.

As a therapeutic unit, SNORS+ has proved extremely useful in the assessment and management of many speech disorders, allowing objective evaluation of the function and coordination of key articulators. It provides visual feedback, which assists in therapy and is very motivating for therapist and patient.

The new parameter addition, electromyographic signal analysis, will allow us to establish the relationship between muscular contraction and resulting movements during speech, as well as new movement measurements, such as lip protrusion, to be assessed.

6. Future work

As has been stated, the initial EMG measurements corresponded to lip protrusion and stretching; future measurements will include mouth opening and closure.

Synchronisation of the EMG data with the SNORS+ system parameters will be necessary for the assessment and correlation between muscular effort and resulting lip movement, during speech rehabilitation sessions.

References:

[1] Borden, G. J., Harris, K. S., and Rafael L. J. "Speech Science Primer", 3rd. Ed. Williams and Wilkins, Baltimore, 1994.

[2] Sharp, P., Kelly, S., Main, A. and Manley, G. "An Instrument for the Multiparameter Assessment of Speech". Medical Engineering & Physics, 1999, 21(9), pp. 661-671.

[3] Moore, Walter H., Jr. "Introduction to Electromyography". Experimental Clinical Phonetics, London, 1984.

[4] Harding, A. and Grunwell, P. "Characteristics of Cleft Palate Speech". European journal of Disorders of Communication, 1996, 31, pp. 331-357.

[5] Tortora, Gerard J. and Grabowski, Sandre R. "Principles of Anatomy and Physiology", 7th. Ed. Harper-Collins, 1992.

[6] Huckabee, Maggie Lee. "A Window into Speech Rehabilitation", 2001, pp. 2.

[7] Cram, Jeffrey R. & Kasman, Glenn S. with Holtz, Jonathan, "Introduction to Surface Electromyography", 1997, pp. 242-243.

[8] DelSys Incorporated. "Surface Electromyography: Detection and Recording", 1996, pp. 1-10.

[9] Gabal, Ahmed A. "Speech Analysis System", 1996, pp. 6-11.