

OPTOPALATOGRAPH (OPG): A NEW APPARATUS FOR SPEECH PRODUCTION ANALYSIS

A. A. Wrench¹, A. D. McIntosh² and W. J. Hardcastle¹

1. Department. of Speech and Language Sciences, Queen Margaret College, Clerwood Terrace, Edinburgh, EH12 8TS UK.

2. Institute for Astronomy, Dept. of Physics and Astronomy, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ UK.

ABSTRACT

This paper describes the early development of a device for measuring tongue-palate distance, contact and pressure across the whole of the hard palate. The Optopalatograph (OPG) is similar in principle to the Glossometer and similar in configuration to the Electropalatograph. It uses optical fibres to relay light to and from the palate and distance sensing is achieved by measuring the amount of light reflected from the surface of the tongue. This paper compares the performance of an optical-fibre-based OPG sensor with that of the direct Glossometer configuration and demonstrates improvements in accuracy and sensitivity.

1. INTRODUCTION

Electropalatography (EPG) was developed more than 30 years ago and over the years its use has extended from the research lab into the clinic where, as a speech therapy tool, it is used in the treatment of a large variety of disorders. Although EPG has been demonstrably of great value, it is restricted in what it can show. If the tongue is not touching the palate there is no indication of how far it is from the palate. Consequently, it is best at providing feedback for obstruent, lateral, and high front vowel articulations. Conversely, when the tongue is in contact with the palate, EPG provides no indication of how firmly the contact is made. Other articulatory measurement apparatus such as the Electromagnetic Articulograph (EMA) and x-ray microbeam can be used to calculate tongue-palate distance but the measurement points are restricted to the midsagittal plane. Furthermore, these techniques are not suited to clinical use due to the preparation time required for collecting the data.

Some work has been carried out to try to extract 3D vocal tract shapes by extending the EPG system. Two similar but independent systems use digitised palate shape data derived from structured light to display the tongue-palate contact pattern in 3D [1][2]. Some examples are given in which fricative data is processed by generating a composite frame for the entire fricative, and computing an area estimate for each row of the composite frame using the assumption of a flat tongue[1]. In another approach [3] the combination of EPG and EMA complement each other; providing an improved representation of the shape and movement of the tongue which reveals details of tongue kinematics not obtainable from EPG alone.

These techniques for enhancing EPG provide some insight into the dynamics of the oral cavity but they require assumptions to be made in order to extrapolate to lateral regions of the oral cavity which are not directly measured. Clinical and scientific research would benefit greatly from an instrument which could provide an indication of the tongue-palate distance as well as contact across the whole palate.

Consideration was given to several technologies which might form the basis of a device with these capabilities. Capacitive and piezoelectric ultrasound techniques were examined but the most promising approach is based on optical distance sensing.

2. OPTICAL DISTANCE SENSING

The technique of measuring reflected light intensity to calculate the tongue/palate distance was first proposed in 1978 by Chuang and Wang [4] and developed throughout the 1980's by Fletcher and McCutcheon [5].

2.1 Principle of Operation

The technique exploits the relationship between the intensity of light reflected from a surface and the distance between the light source/sensor and the reflecting surface. A light beam radiating from a point source strikes the diffusive surface of the tongue (Figure 1.). A photosensor directed at the surface of the tongue collects a proportion of the light scattered in its direction. The intensity of light at the surface is inversely proportional to the

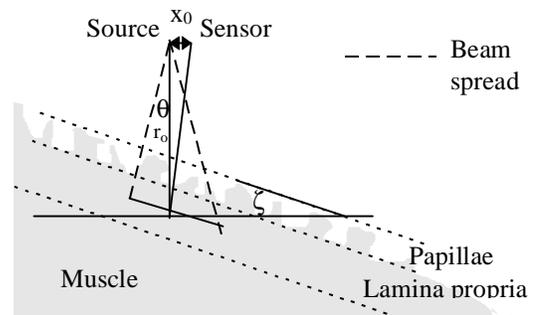


Figure 1. Factors affecting light reflected from the translucent tongue tissue.

square of the distance to the source r_0 as $1/4 \pi r_0^2$. The intensity of light collected at the sensor depends on three factors:

- a) The angle of divergence of the beam θ
- b) The angle of inclination ζ that the beam makes with the surface
- c) The distance between the source and the sensor x_0

The first factor is a constant for a given sensor system and can be accounted for as part of the calibration procedure. The second factor is more important and was shown by Chuang and Wang [4] to lead to an 8% overestimation for an inclination angle of less than 15° . We will examine this factor in greater detail in the Performance section of this paper. The latter factor is not significant if the tongue-palate distance r_0 is greater than the source/sensor separation x_0 . This condition is always true in practice because the translucent nature of the tongue tissue permits light to be reflected from structures below the surface of the tongue. For this reason, even if the sensor is pressed hard into the tongue the light intensity continues to rise. Indeed, this property provides a means for determining the firmness of contact between tongue and palate.

2.2 Glossometer

Fletcher and McCutcheon modified the technology and procedures to make the reflected light sensing device more practicable for experimental and clinical use and gave it the name Glossometer. The artificial palate is thin (0.3mm) but the sensors are $2 \times 3 \times 6$ mm and the palate can reach 5mm in thickness to embed them securely. Consequently the sensors are placed posterior to the alveolar ridge and typically 4 are located along the midline of the artificial palate. A calibration technique was introduced consisting of a set of spacer grids which clip to the teeth and maintain the tongue at a known distance from the palate. The calibration data is stored in a lookup table on the computer and this mapping accounts for variations in reflectivity of the tongue and surrounding structures such as the teeth.

Fletcher et al have used the Glossometer to teach vowels to profoundly hearing-impaired speakers with some success [6]. However, the device has not found wide acceptance in either the scientific or clinical fields. The midsagittal arrangement of sensors limits the data which can be recorded by the Glossometer and offers no scientific advantage over that which is provided by EMA and x-ray microbeam. In addition, the size of the artificial palate creates problems for the wearer when articulations with narrow tongue palate constrictions are attempted and is consequently best used for the examination of vowels. In its favour, the system is considerably cheaper than either EMA or microbeam and is more practicable for clinical use.

3. OPTOPALATOGRAPH

The Optopalatograph is designed to combine the benefits of Electropalatography and Glossometry to produce a device capable

of measuring the volume and shape of the intra-oral cavities and constrictions.

The key difference between the Glossometer and the Optopalatograph is the removal of the light emitters and receivers from the palate and instead transmitting the light by optical fibre to a remote distance sensing unit. This design change removes two of the major obstacles to the wider acceptance of a light distance sensing system as a viable clinical and scientific instrument. Firstly, the use of fibres results in a less cumbersome palate construction which enables many more measurement points to be provided. Crucially, this allows measurement points off the midline thus providing data which is unavailable from EMA or microbeam apparatus. Secondly, the palate can be manufactured with a uniform thickness of 1-2mm compared with 3-5mm that results from embedding the Glossometer sensor. This reduces the articulatory adjustment required when the palate is worn, particularly for coronal obstruents. The benefits provided by the Optopalatograph may be summarised as follows:

- It permits a *maximum* of 70 measurement points.
- The measurement points can be spread across the whole palatal surface
- The optopalate is no thicker than a conventional electropalate (0.5-2mm)
- Receiver and transmitter fibres can be positioned closer together reducing the effect of the angle of the reflecting surface (the tongue).
- The optopalate has no electrical connections leading into the mouth.
- Larger, more powerful remote emitters can be used, reducing sensitivity to noise and increasing the range.

3.1. Optical Fibres

The key element in the design of the Optopalatograph is the use of optical fibres. The optical fibres must meet physical as well as optical performance criteria. They must be robust to handling; have a small bend radius, ideally permitting a 90° bend with 2mm radius so that light shines perpendicular to the palate surface; and they must have a diameter small enough that the accumulation of fibres exiting the mouth is not too great. A reasonable upper limit on the diameter of fibre bundles exiting the mouth is 2.5mm.

Glass fibres are readily available at diameters as low as $50\mu\text{m}$ but were found to be too brittle. The requirements are however satisfied by 0.25-0.5mm plastic optical fibre. Allowing for two 2.5mm bundles, one at each side of the mouth, the device is constrained to a maximum of 38 0.5mm fibres or 140 0.25mm fibres or an intermediate number if mixed fibre diameters are used. Plastic optical fibres exhibit transmission attenuation better than 3dB for wavelengths in the visible to near infrared range (300-880nm). A more common source of attenuation is from

fractures due to rough handling or abrupt bends. Although there is a degree of light lost, this range of plastic fibres have been found to have adequate transmission with bend radii of 2mm and protective sleeving can be used to reduce handling damage..

The smaller the fibre diameter the lower the light levels which must be transduced and amplified, resulting in increased noise levels both from ambient light sources and electrical sources. The choice between 0.25 and 0.5mm fibre diameter is therefore a trade off between the number of measurement points and the signal to noise ratio of the system.

3.2. Sensors

High intensity light emitting diode (LED) sources are used. LEDs are preferred because they are easily switched, compact and have a stable and reliable level of light emission. High power 3.5cd red or 6cd amber AALGAS LEDs may be used to minimise ambient light interference.

IR transmitters can reach higher output levels if they are pulsed. at high frequencies (10kHz), improving the signal to noise ratio by as much as 10dB. If this pulsed signal is also decoded with a Phase-locked Loop then it is possible to differentiate the ambient radiation from the signal and provide an indication that the noise level is masking the signal. However, precautions are still necessary to minimise the ambient light level entering the oral cavity.

To eliminate interference from other emitters, each emitter is switched on in sequence and in synchrony with the associated sensor amplifier. In future systems, surrounding sensors may also be activated and the levels used to compensate for distance estimation errors due to the angle of inclination of the tongue surface.

4. PERFORMANCE

The introduction of optical fibres as a medium for transmission of light to and from a remote photoemitter/sensor array require a set of operational parameters which are quite different from the Glossometer. Intensity of light emerging from the optical fibre (I),

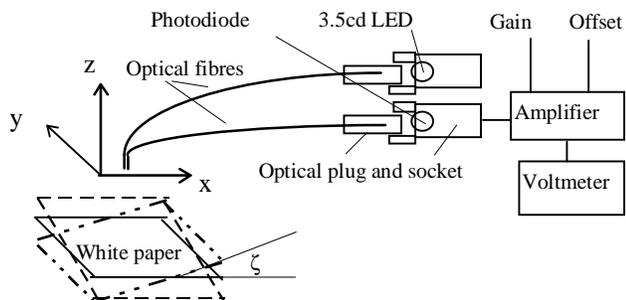


Figure 2 Experimental apparatus for evaluating the performance of an OPG sensor

beam spread (θ) and proximity of transmitting and receiving fibres (x_0) are constrained by physical factors unique to the OPG.

The LED is not limited in size and high power 5mm LEDs can be used but, in the absence of a collimating system, the diameter of the fibre limits the amount of light reaching the palate. The largest practical fibre size is 0.5mm. The performance of the OPG was tested using a prototype system consisting of a 3.5cd red ALGAAS emitter 2x 400mm lengths of 0.5mm diameter plastic optical fibre and a photodiode connected to a high input impedance linear amplifier with adjustable gain and offset.

Beam spread is dependent on the angle that the light enters the fibre and was found to be $\theta = 20^\circ$ for the prototype. The ends of the fibre were however unpolished. It may be possible in the future to use lenses to focus the light as a parallel beam entering the fibre thus reducing beam spread on exit.

Figure 3. shows the signal output vs. distance using a white blotting paper and pork muscle as reflecting surfaces. The results show a usable range up to 40mm. There is reason to expect improvements to this performance if a collimating system is used

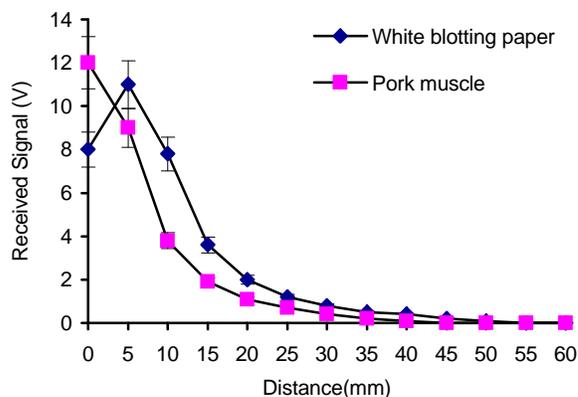


Figure 3 Experimental distance sensing functions for reflected light on white blotting paper and muscle tissue

to concentrate the light level entering the fibre. Accuracy of the system depends on errors due to angle of inclination of the reflecting surface. The translucent pork muscle is used to simulate the tongue and, as predicted and in contrast to the blotting paper shows no reduction in received signal level at close range.

Figures 4 and 5 show the variation in signal response as the reflecting surface is inclined at angles increasing by 10° intervals up to $\pm 40^\circ$ with reference to the y axis and x axis respectively. The transmitting and receiving fibres are placed in parallel, in the x/z plane and touching so that the separation x_0 is 0.5mm centre to centre. These graphs can be interpreted as percentage error in distance using the response curve for white blotting paper in figure 3. In figure 4. when the axis of rotation of the reflecting

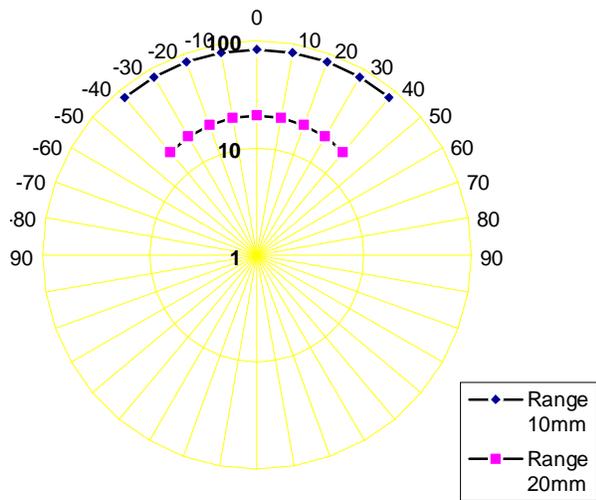


Figure 4 The influence of parallel rotation of the reflector on the sensor output; rotating axis parallel to the light source/sensor axis (c.f. Chuang and Wang Figure 4.[4])

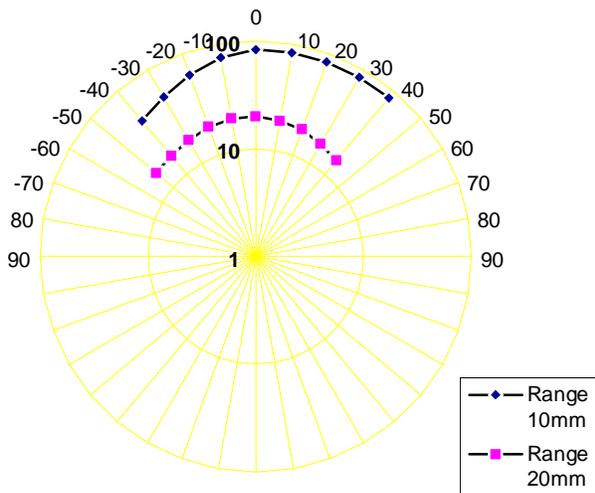


Figure 5 The influence of perpendicular rotation of the reflector on the sensor output; rotating axis perpendicular to the light source/sensor axis (c.f. Chuang and Wang Figure 5.[4])

surface lies on the x/z plane in which the fibres are aligned then the distance is estimated accurately up to 25° and is overestimated by 10% at 40° . In figure 5 when the axis of rotation of the lies perpendicular to the x/z plane in which the fibres are aligned then the distance is overestimated by 10% at $\zeta = +20^\circ$ and 5% at -20° .

4. CONCLUSION

The data provided in this paper shows that the results of remote optical distance sensing using optical fibres are comparable to those of Chuang and Wang who measured reflected light using emitters and transducers mounted directly on the palate. The close proximity of the transmitting and receiving fibres reduces the error in distance estimation due to angle of incidence. This means that the OPG is less sensitive to sensor orientation. The system is accurate to within 10% regardless of sensor orientation for angles of inclination up to 20° . i.e. $\pm 0.5\text{mm}$ at a range of 5mm and $\pm 2\text{mm}$ at a range of 20mm.

Further research using a full palate based prototype is underway to determine the accuracy of the system but, based on the current comparison, the system is expected to perform at least as well as the Glossometer. However, with the larger number and wider distribution of measurement points across the entire palate, the OPG will provide 3D data on oral cavity configurations unavailable from other instrumentation.

The Optopalatograph measures relative tongue position with reference to the palate surface. This makes it compatible with gestural analyses where the vocal tract parameters relate to constriction size rather than absolute position of the articulators. It is envisaged that the OPG may be used in conjunction with other more accurate measurement systems such as EMA whose data could be used to verify and calibrate the OPG data. It may also be used alone as a clinical tool for providing visual feedback in speech therapy.

5. REFERENCES

1. Chiu, W.S.C., Shadle, C. H. and Carter, J.N., "Quantitative measures of the palate using enhanced electropalatography" *European Journal of Disorders of Communication*, Vol. 30, pp 149-160, 1995.
2. Saw, C. C., "Customized 3-D Electropalatography display." ASA 127th Meeting, 1pSP28. M.I.T. June 6-10, 1994
3. Hoole, P., Nguyen-Trong, N., and Hardcastle, W., "A comparative investigation of coarticulation in fricatives: electropalatographic, electromagnetic and acoustic data." *Language and Speech*, Vol. 36, pp 235-260, 1993.
4. Chuang, C-K and Wang, W, S-Y, "Use of optical distance sensing to track tongue motion", *Journal of Speech and Hearing Research*, Vol 21, pp 482-496, Sept., 1978.
5. Fletcher, S.G., McCutcheon, M.J., Smith, S.C. and Smith W.H. "Glossometric measurements in vowel production and modification," *Clinical Linguistics and Phonetics*, Vol. 3, No. 4, pp 359-375, 1989.
6. Fletcher, S. G., Dagenais, P. A., Critz-Crosby, P., "Teaching Vowels to Profoundly Hearing-Impaired Speakers Using Glossometry", *Journal of Speech and Hearing Research*, Vol. 34, pp 943-956, 1991.