A NEW ARTIFICIAL PALATE DESIGN FOR THE OPTICAL MEASUREMENT OF TONGUE AND LIP MOVEMENTS

Peter Birkholz and Christiane Neuschaefer-Rube

Clinic of Phoniatrics, Pedaudiology, and Communication Disorders
University Hospital Aachen and RWTH Aachen University
peterbirkholz@gmx.de

Abstract: Electropalatography (EPG) is a well-known and established technique to measure the detailed time-varying contact pattern between the tongue and the hard palate. For EPG, the speaker wears an artificial palate in his mouth, which contains multiple electrodes (contact sensors) to register contact with the tongue. A related technique, albeit less known, is glossometry or optopalatography (OPG). For OPG, the speaker wears an artificial palate that contains optical distance sensors instead of contact sensors. These sensors allow to measure the distance between the palate and the tongue at multiple points and so to reconstruct part of the tongue shape, regardless of whether the tongue is touching the palate or not. Despite this advantage for the analysis of speech movements compared to EPG, the technique was not widely adopted. One of the main reasons is the high costs of manufacturing the corresponding palates, which must be individually made for each speaker. This study presents a new palate design for OPG that greatly reduces the cost and time needed to make the palates. Furthermore, the new design allows the additional measurement of the upper lip position. The effectiveness of the palate design is demonstrated by measurements of vocalic tongue shapes and lip positions.

1 Introduction

Electropalatography (EPG) is a well-known technique to measure and visualize the time-varying pattern of contact between the tongue and the hard palate [6]. It is widely used in many phonetic laboratories and speech therapy clinics throughout the world. For EPG, the speaker wears a custom-made artificial plate molded to fit his hard palate with a number of electrodes mounted on its surface to detect lingual contact. Compared to other techniques to measure articulatory movements, e.g., cineradiography, x-ray microbeam, ultrasound, electromagnetic articulography, or real-time magnetic resonance imaging (see [3] for details), the main advantages of EPG are low device cost and ease of use. Once an artificial palate (pseudopalate) is made for a speaker, it can be re-used by the speaker as often as needed with only little time for preparation and acclimatization. A drawback of EPG is that it provides no indication about the position of the tongue when it is not touching the palate. Therefore, it is not suitable for the analysis of phones with little or no tongue-palate contact, for example mid and open vowels.

A technique closely related to EPG is optopalatography (OPG). It uses optical distance sensors mounted in a pseudopalate to transduce the distance between the palate and the tongue at multiple measurement points. A distance sensor consists of a light emitting diode (LED) and a photosensor. The LED radiates a light beam towards the tongue surface, and the photosensor collects the proportion of light that is scattered back in its direction. Because the intensity of light detected by the photosensor decreases monotonically with increasing distance, the distance can be inferred from the photosensor measurement. This kind of distance sensing in the mouth was first used in 1978 [2] and refined by Fletcher et al. in the 1980s [5, 4]. However, because of a number of limitations, the technique was not widely adopted. The major drawback was the size of the sensors, which made the pseudopalates rather thick (up to 4 mm [5]) and therefore inconvenient to speak with. Furthermore it was time-consuming and therefore expensive to make the palates. Wrench et al. [8, 9, 10] used optical fibres to transmit the emitted and reflected light between the measurement points on the pseudopalate and a distance-sensing unit outside the mouth, which contained the LEDs and photosensors. This allowed to manufacture the palates with a uniform thickness of only 1-2 mm. However, the difficulties in making palates with optical fibres as well as calibration issues prevented the widespread use of this innovation as well. Recently, Birkholz et al. [1] created a pseudopalate that contained both contact sensors as for conventional EPG and optical distance sensors as for OPG. However, the arrangement of the optical sensors as well as the used sensor type were not optimal, and making the pseudopalate was rather difficult.

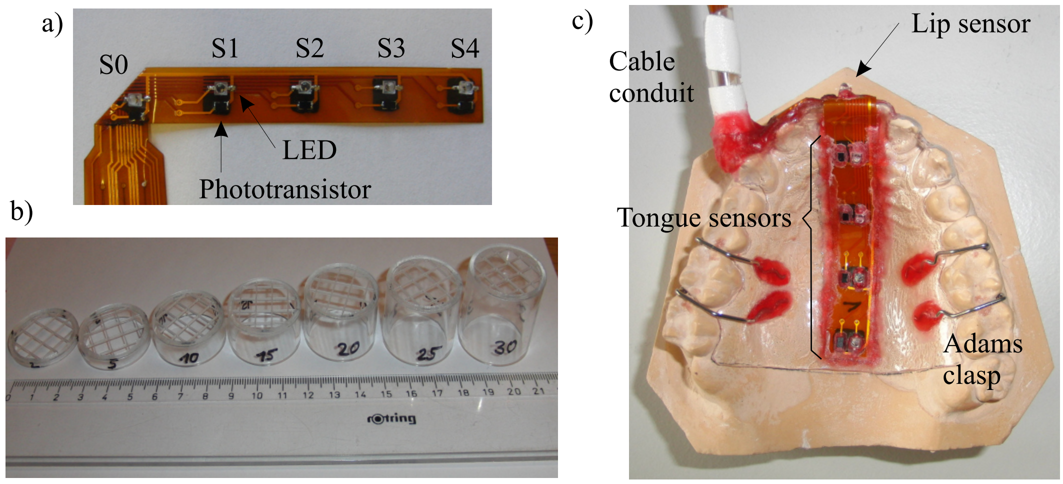

Figure 1 - a) Flexible circuit for a pseudopalate populated with the optical distance sensors S1-S4 (LED-phototransistor pairs). b) Selected spacers used to obtain the distance sensing function of the individual sensors. c) Finished pseudopalate mounted on a plaster model of the hard palate.

In this paper, we present an OPG palate design that addresses the above issues. Our design allows to make a pseudopalate in only 2-3 hours with low material costs and rather small distance sensors based on precast flexible circuits in a defined range of sizes. Compared to previous OPG palates, the proposed design contains an additional sensor to measure the position of the upper lip. Furthermore, the sensors can be conveniently calibrated in-vitro. The proposed palate design is evaluated with respect to the analysis of tongue shapes and lip positions for vowels.

2 The new palate design

The proposed palate design is shown in Fig. 1c. The key features of this design are

- the use of a single flexible circuit board that carries the distance sensors - both the sensors to measure the tongue-palate distance and the upper lip position (Fig. 1a). The use of flexible circuit strips proved to be very effective for the manufacturing of conventional EPG palates [7] and was therefore adopted for the present OPG-palate design. Using a flexible circuit board not only reduces the time to make a pseudopalate, but also the costs, when the circuit boards are manufactured in a limited range of sizes (Sec. 2.2).

- the use of *discrete* LED-phototransistor pairs for the distance sensors. Using discrete LEDs and phototransistors instead of combined sensor units as in previous designs (e.g. [2, 5, 1]) allowed to create tailored sensors that were small in size, had a sufficient measuring range, and were rather resistant to measurement errors caused by saliva coating. Commercially available combined sensor units were usually not designed with respect to these criteria.
- a high wearing comfort. The base layer of material for the pseudopalate consists of a thin 0.5 mm foil thermoformed over a plaster model of the hard palate. The foil is removed from the molars and premolars to allow to completely close the jaw and to prevent saliva to pool between the teeth and the pseudopalate. Around the incisors and canines, the foil is preserved to retain the palate and allow to fix the anterior part of the flexible circuit with the lip sensor. To retain the pseudopalate in the posterior part, Adams clasps were fixed on the foil with modeling resin.
- easy in-vitro sensor calibration as described in Sec. 2.1.

Hence, making a pseudopalate just requires a few steps: thermoforming the foil over the plaster model of the palate, fixing the flexible circuit with the sensors, removing the foil from the molars and premolars, adapting and fixing the Adams clasps, and fixing a silicone cable conduit to lead the flexible circuit outside the mouth.

2.1 Optical distance sensors

As stated above, each distance sensor consists of a discrete LED and phototransistor soldered on the flexible circuit board. Testing different combinations of available LEDs and phototransistors for infrared light with respect to measuring range and size, we found the LED VSMY2850 and the phototransistor TEMT7100 (both by Vishay Semiconductors) with a gap of 0.6 mm between them to be a suitable compromise. The LED contains a lens and has a footprint of 2.3 x 2.3 mm^2 and a height of 2.8 mm. The phototransistor has no lens and a tiny 0805 package with height of only 0.8 mm. After soldering the components on the flexible circuit board, the electrical contacts and margins were sealed with modeling resin to prevent contact with saliva. The optical properties of the sensors are usually not perfectly identical, because of small placement variations of the components after manual soldering, and because the modeling resin around the components slightly influences their optical characteristics. Therefore, each sensor was calibrated individually. For the calibration of a sensor, its “distance sensing function”, i.e., the output of the phototransistor (after analog-to-digital conversion) for a range of distances to the tongue surface, was measured. Therefore, the flexible circuit was temporarily fixed on a portable flat surface as in Fig. 1a, and using custom-made spacers, the tongue surface was brought into different distances from the sensors as illustrated in Fig. 2a. The spacers (Fig. 1b) consisted of plexiglas tubes (26 mm inner diameter; lengths 2, 5, 10, 15, 20, 25, 30 mm) with a wide-meshed strong net spanned over the opening for the protruded tongue. For each sensor, the phototransistor output was measured as a function of the distance. The linearly interpolated distance sensing functions for the five sensors of one of the prototypes are shown in Fig. 2b. These functions were later directly used to map a given sensor output to a distance value.

During the measurements, the distance sensors were switched in sequence to avoid optical cross-talk. The output of a sensor was taken as the 10-bit ADC value of the voltage between

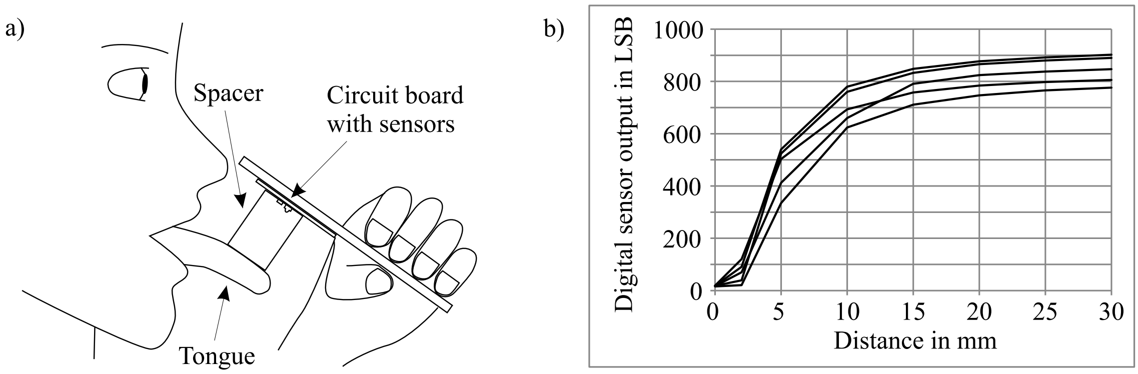

Figure 2 - a) The way we put the tongue in different distances from the sensors using the spacers. b) Distance sensing functions for the five optical distance sensors after soldering them on the flexible circuit and sealing them with modeling resin.

a high-side resistor and the phototransistor connected in series between the supply voltage and ground. Each sensor was sampled at a rate of 100 Hz with each sample corresponding to the average of four consecutive measurements.

2.2 Standard sizes for flexible circuits

The single-unit production of flexible circuits is rather costly. But the price per unit drops significantly when multiple identical circuit boards are produced, for example in a small range of sizes that fit the majority of people. To obtain such a set of sizes, we made a survey of the 3D palatal shape of a representative group of people.

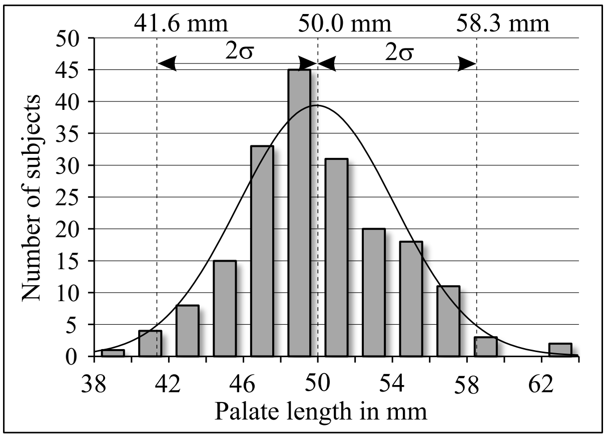

Figure 3 - Distribution of the palate length of of 191 people measured along the palatal arch between the Ah-line and the posterior onset of the upper incisors.

We analyzed the palates of 85 male and 106 female individuals with ages between 9 and 55 years (mean: 19.4 years; s.d.: 6.96 years). For each individual, a plaster model of the hard palate was scanned with a high-resolution 3D-Scanner (3shape D700 by Bego) to obtain a 3D-digital wireframe-model. Using the Software OrthoAnalyzer^{TM}, we measured the length of each palate along the palatal arch between the Ah-line and the posterior onset of the upper incisors, as well as the posterior and anterior height of the upper incisors. These three measurements determine the required length for a flexible circuit strip and the distribution of the optical sensors along the strip. The histogram of the palatal arch length is shown in Fig. 3. This distribution has a mean value of 50.0 mm and a standard deviation σ of 4.2 mm. To cover approximately the $\pm 2\sigma$-range of the distribution, i.e., 95% of the data, five different lengths of the flexible circuit for the palatal arch (42 mm, 46 mm, 50 mm, 54 mm, 58 mm) would be appropriate for a maximal fitting error of 2 mm. Fewer sizes could be used if a higher error is tolerated. The length of the upper incisors was on average 8.5 ± 1.2 mm (mean±s.d.) at the posterior side and 10.6 ± 1.2 mm at the anterior side. For the present study, multiple equal flexible circuits were designed and produced only for the average size (50 mm for the palatal arch length) to make pseudopalates for two people.

3 Evaluation

3.1 Material and data processing

To demonstrate the potential of the proposed system, the tongue shapes and lip positions were measured for the German vowels /a, e, i, o, u, ɛ, ø, y, ə/ and compared with respect to the common articulatory features height, backness, and roundedness. For each of the eight non-neutral vowels, ten German words with the respective vowel in the nucleus of the stressed syllable were recorded from one speaker (PB) wearing the pseudopalate. For the vowel /i/, for example, the words *Biere*, *Tiere*, *Kino*, *Miete*, *niesen*, *siegen*, *schieben*, *Riegel*, *Ziegel*, and *Zwiebel* were spoken. Hence, the vowels were produced in a variety of different contexts. The audio signal was recorded synchronously with the articulatory pseudopalate data at a sampling rate of 22050 Hz using a standard headset. The neutral vowel /ə/ was produced in isolation for several seconds and recorded accordingly.

At the acoustic midpoint of each target vowel in the words, the measurements of the optical distance sensors were used to reconstruct the contour of the tongue and estimate the position of the upper lip. The 2D mid-sagittal location of the sensors and their optical axes were estimated from a photograph of a side view of the pseudopalate.

3.2 Results and discussion

Figure 4 shows the average tongue contours and lip positions for the considered vowels. The gray regions indicate the $\pm 1\sigma$ range of the 10 samples per vowel. The tongue shape and lip position of the neutral vowel /ə/ are drawn with dotted lines as a reference in all subimages. The subimages for the vowels were arranged with respect to the articulatory features height and backness similar to the vowel chart of the International Phonetic Alphabet (IPA). This shows the close correspondence between the phonetic features and the measured tongue shapes. Also note the small variance of the tongue shapes of the 10 samples per vowel, which indicates a high repeatability of the measurements.

The measured lip positions need some explanations. They do not directly reflect the degree of lip protrusion or rounding. Instead, a high distance value was measured for the unrounded vowel /i/ and a small value was measured for the rounded vowel /y/, for example. This is because for unrounded vowels, the upper lip covers the lip sensor less than for rounded vowels. In fact, for /i/, /e/, or /a/, the lip was mostly above the optical axis of the lip sensor so that only little light of the LED was reflected back to the phototransistor, corresponding to a high distance. In contrast, for the rounded vowels /u/ or /y/, the upper lip was closely in front of the incisors, corresponding to a short distance. Therefore, the “lip positions” measured with the

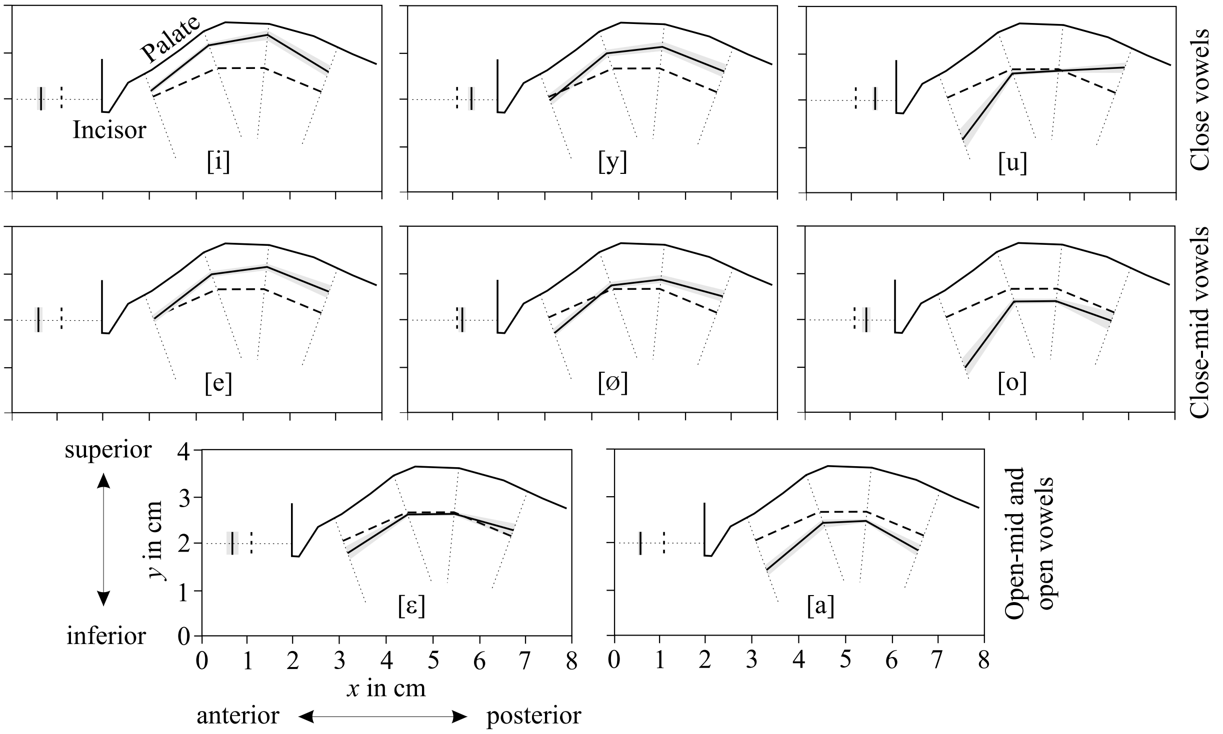

Figure 4 - Measured tongue shapes and lip positions of the vowels. The gray regions indicate the $\pm 1\sigma$ range of the samples. The dotted lines represent the tongue shape and lip position for the neutral vowel /ə/ as reference. The thin dotted lines indicate the optical axes of the distance sensors.

pseudopalate are rather inversely related to the degree of lip rounding.

4 Conclusion

We presented a new palate design for the optical measurement of the anterior tongue contour and the upper lip position. The design was optimized with respect to short time and low cost for manufacturing the palates as well as high wearing comfort and measurment accuracy. A skilled dental technican needs about two hours to make a new palate given a populated flexible circuit board. The material costs are about 50 Euros per pseudopalate when the flexible circuit boards are ordered in volumes of 5-10 units. For our current prototypes, four sensors were used to reconstruct the shape of the tongue. However, there is room for up to six or seven sensors along the midline for a more detailed reconstruction of the tongue. The current calibration method using the tongue spacers works well, but a fully automatic calibration method without the need for a “real” tongue is under consideration.

5 Acknowledgments

We would like to thank Martin Humperdinck, Philippe Dächert, Alex Röll, Alexander Füglein, Ulrike Fritz, André Barloi, Udo Höhn, Sebastian Scharrmann, René Bohne, and Ingmar Steiner for their valuable contributions to this study. The study was funded by the German Research Foundation, grant no. BI 1639/1-1.

References

[1] BIRKHOLZ, P. and C. NEUSCHAEFER-RUBE: *Combined optical distance sensing and electropalatography to measure articulation.* In *Interspeech 2011*, pp. 285–288, Florence, Italy, 2011.

[2] CHUANG, C.-K. and W. S. WANG: *Use of optical distance sensing to track tongue motion.* Journal of Speech and Hearing Research, 21:482–496, 1978.

[3] EARNEST, M. M. and L. MAX: *En Route to the three-dimensional registration and analysis of speech movements: instrumental techniques for the study of articulatory kinematics.* Contemporary Issues in Communication Science and Disorders, 30:2–25, 2003.

[4] FLEGE, J. E., S. G. FLETCHER, M. J. MCCUTCHEON and S. C. SMITH: *The physiological specification of American English vowels.* Language and Speech, 29(4):361–388, 1986.

[5] FLETCHER, S. G., M. J. MCCUTCHEON, S. C. SMITH and W. H. SMITH: *Glossometric measurements in vowel production and modification.* Clinical Linguistics and Phonetics, 3(4):359–375, 1989.

[6] HARDCASTLE, W., W. JONES, C. KNIGHT, A. TRUDGEON and G. CALDER: *New developments in electropalatography: A state-of-the-art report.* Clinical Linguistics and Phonetics, 3(1):1–38, 1989.

[7] WRENCH, A. A.: *Advances in EPG palate design.* Advances in Speech-Language Pathology, 9(1):3–12, 2007.

[8] WRENCH, A. A., A. D. MCINTOSH and W. J. HARDCASTLE: *Optopalatograph (OPG): A new apparatus for speech production analysis.* In *4th International Conference on Spoken Language Processing (ICSLP 1996)*, pp. 1589–1592, Philadelphia, PA, USA, 1996.

[9] WRENCH, A. A., A. D. MCINTOSH and W. J. HARDCASTLE: *Optopalatograph: Development of a device for measuring tongue movement in 3D.* In *EUROSPEECH '97*, pp. 1055–1058, Rhodes, Greece, 1997.

[10] WRENCH, A. A., A. D. MCINTOSH, C. WATSON and W. J. HARDCASTLE: *Optopalatograph: Real-time feedback of tongue movement in 3D.* In *5th International Conference on Spoken Language Processing (ICSLP 1998)*, Sydney, Australia, 1998.