INTERSPEECH 2011

Combined optical distance sensing and electropalatography to measure articulation

Peter Birkholz, Christiane Neuschaefer-Rube

Clinic for Phoniatrics, Pedaudiology, and Communication Disorders
University Hospital Aachen and RWTH Aachen University
Pauwelsstrasse 30, 52074 Aachen, Germany
pbirkholz@ukaachen.de, cneuschaefer@ukaachen.de

10.21437/Interspeech.2011-111

Abstract

We present the first prototype of a new optoelectronic instrument for the combined real-time measurement of the tongue contour in the mid-sagittal plane, the contact pattern between the tongue and the palate, and the position of the lips. The instrument consists of a thin acrylic pseudopalate with embedded contact sensors, as for electropalatography, and optical distance sensors to measure tongue-palate distances, as for glossometry. One additional distance sensor is located at the anterior side of the upper incisors to register the degree of opening and protrusion of the lips. Together, the sensors provide complementary information about the articulation of vowels and consonants, which was verified in initial experiments. The instrument offers new perspectives for the study of normal and disordered speech production, as well as for silent speech interfaces and speech prostheses for laryngectomees.

Index Terms: electropalatography, optopalatography, glossometry, optical distance sensing

1. Introduction

The instrumental analysis of articulation is indispensable for advances in the understanding of normal and disordered speech production. A highly effective and rather cheap technique for the analysis of spatio-temporal contact patterns between the tongue and the palate is electropalatography (EPG) [1]. For this technique, the user wears an artificial plate (pseudopalate) moulded to fit the upper palate with a number of contact sensors distributed over its surface to detect lingual contact. It provides valuable data for those sounds that involve considerable lingual-palatal contact, like many consonants, high vowels, and diphthongs. However, to track the motion of the tongue for low or back vowels, or the motion of the lips or the velum, other techniques must be used, e.g. ultrasound, electromagnetic articulography, or X-ray microbeam [2]. These techniques are typically more expensive or complex than EPG.

However, already in 1978, Chuang and Wang [3] demonstrated that also the tongue contour in the frontal oral cavity can be measured using a pseudopalate. They introduced a device that contained four optical distance sensors along the midline of a pseudopalate (instead of contact sensors as in EPG palates) to determine the distance of the tongue from the palate. Fletcher et al. [4] advanced this technique in the 1980th and gave it the name glossometry, but apart from a few clinical studies, it gained no widespread use.

Later, Wrench et al. [5, 6] picked up the idea of optical distance sensing to measure the tongue shape and position based on a pseudopalate and developed the optopalatograph. This device uses pairs of optical fibres to transmit light between the measurement points on the pseudopalate and a remote distance sensing unit. This way, the pseudopalate could be made thinner than the pseudopalates for glossometry to improve the comfort for the user. Furthermore, due to the little space required for the optical fibres, more measurement points could be provided, especially off the midline. However, optical fibres are very hard to work with and making handmade palates of this type is therefore expensive.

In this study, we present the prototype of a pseudopalate that combines optical distance sensing with classical EPG. Together, this should provide complementary real-time data about the articulation of many consonants in terms of tongue-palate contact patterns and the articulation of vowels in terms of tongue contours. Besides the EPG contact sensors, the palate contains three distance sensors to measure the tongue-palate distance. Aditionally, we embedded one forward-directed sensor in front of the upper incisors to measure the state of the lips, and one rear-directed sensor at the posterior end of the palate to measure the state of the velum. This paper reports about the design and manufacture issues of the pseudopalate, introduces the measuring system, and presents the results of initial performance tests of the system.

2. Choice of the distance sensor type

Optical (reflective) distance sensors consist of an infrared light emitter, typically a light emitting diode (LED), and a light detector, typically a photodiode or a phototransistor. The emitter illuminates the surface in front of the sensor and the detector measures the amount of reflected light, which reduces with increasing distance. This allows to infer the distance from the reflected light level. Virtually hundreds of optical distance sensor types are available today for different applications. For our application, the sensors must satisfy an number of specific requirements that are detailed in the following.

The major requirements are small size and high operating range. Only small sensors will leave enough space on the pseudopalates for the regular EPG contact sensors and make the palate thin enough to be comfortable. The operating range should be at least 25 mm, i.e. a palate-tongue distance of 25 mm should be detectable with acceptable precision. Greater distances rarely occur in continuous speech [3]. Further selection criteria are the good availability of the sensors (no obsolete or discontinued parts), the beam angle of the light source (a small beam angle for a more focussed illumination is preferable), and the distance between the emitter and detector on the sensor. A small emitter-detector distance increases the risk for

28 – 31 August 2011, Florence, Italy

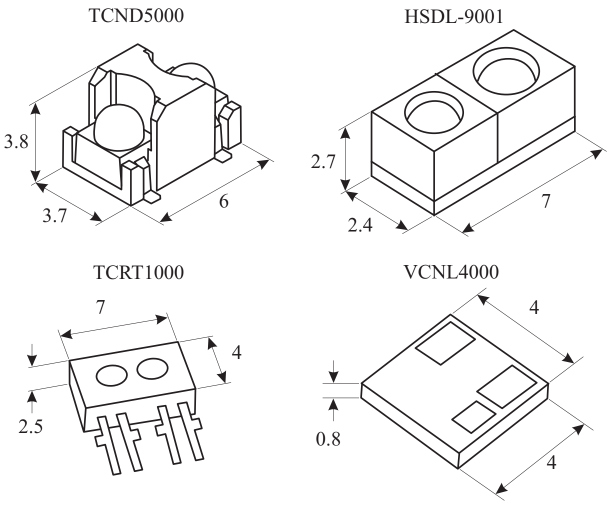

Figure 1: *Selection of potentially suitable distance sensor types. All dimensions are in mm.*

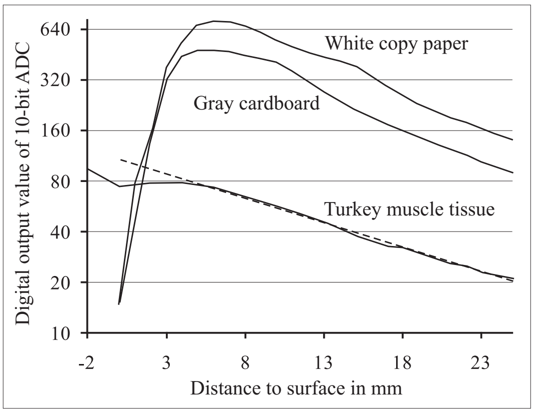

Figure 2: *TCND5000 sensor output as a function of the distance for three different reflectors.*

optical cross-talk, when the sensors become coated with saliva, as reported by Fletcher et al. for their sensors [4]. In this case, light from the emitter is directly transmitted to the detector through the saliva bridge and greatly distorts the distance measurements. On the other hand, the emitter-detector distance may not be to high, because the sensor may otherwise fail to measure distances in the very near range (see below).

We have conducted a comprehensive search among the currently available sensors according to the requirements and eventually found four potentially suitable types. The design and major dimensions of the sensors are shown in Fig. 1, and additional properties are listed in Tab. 1. With respect to size – especially thickness – and operating range, the VCNL4000 would be the first choice. Unfortunately, this sensor was announced not until Dec. 2010 and was not yet available at the beginning of 2011. The TCRT1000 is similar in size and performance to the no longer available sensor SFH-900 that was used by Fletcher et al. [4]. This sensor type showed the same sensitivity to saliva coating as reported by Fletcher et al. [4] in our own pre-tests and was therefore not used for our prototype. The TCND5000 and the HSDL-9001 are sensors with discrete semi-spherical lenses for the emitter and detector and have a very good operating range. Both sensor types seemed equally suited for our application. We decided to use the TCND5000 for the first prototype. The height of these sensors was reduced to less than 3 mm by sanding a bit of material from the bottom of each sensor and the top of the cross-talk barrier.

Table 1: *Properties of distance sensor types.*

Type	Manu-fac-turer	Range for output 20%	Emitter-detector distance
TCND5000	Vishay	25 mm	3.9 mm
HSDL-9100	Avago	25 mm	3.9 mm
TCRT1000	Vishay	4 mm	2.5
VCNL4000	Vishay	high	2.5

Figure 2 shows the output of the TCND5000 as a function of the distance for three types of reflecting surfaces. For white copy paper and gray cardboard as reflector, the curves have a turnaround point at a distance of about 6 mm, i.e. the output level rises from 0 to 6 mm, and then decreases with increasing distance. The cause for this turnaround point is the finite distance between the emitter and the detector of the sensor [3]. This would actually make it impossible to determine a unique distance value from a given sensor output value, because this requires the curve to be strictly monotonic decreasing. Fortunately, the translucent tissue of the tongue prevents a decrease of the reflected light intensity at close ranges, making the method practical at all [3]. Turkey muscle tissue is somewhat similar to tongue tissue, and the corresponding measured curve is shown in Fig. 2. It still has a rather flat region between 0 and 3 mm, but our experiments with a real tongue showed no disambiguities between sensor values and distances at close ranges.

3. Making of the pseudopalate

The design of the pseudopalate is based on the Articulate EPG palate type [7]. The Articulate palate is made up by pressure forming a 0.5 mm sheet of acrylic on a plaster model of the subjects palate, placing flexible circuit strips with the contact sensors on this surface, and then pressure forming a second layer of 0.5 mm acrylic on top to seal the circuits. For our prototype, we additionally glued the five distance sensors on the surface of the first sheet of acrylic according to Fig. 3. One sensor in front of the upper incisors was directed towards the lips, one sensor at the posterior end was directed towards the velum, and three sensors were directed towards the tongue surface. All three distance sensor directed towards the tongue were placed between the fourth and eighth row of contact sensors to retain a high density of contact sensors in the alveolar part of the palate. After the flexible circuit strips with the contact sensors and the distance sensors were fixed to the first sheet of acrylic, the second sheet was pressure formed on top, sealing the sensors. However, with the second acrylic sheet pressed on the distance sensors, we found that the operating range of the sensors was strongly reduced – probably because the light emitted by the LED was now mainly dispersed in the transparent sheet. Therefore, the second sheet was cut out around the distance sensors so that the lenses of the sensor were directly exposed to the mouth cavity. Making a classical EPG palate of the Articulate type without the distance sensors takes about 3-4 hours for a trained dental technician. The preparation and inclusion of the distance sensors took about three additional hours.

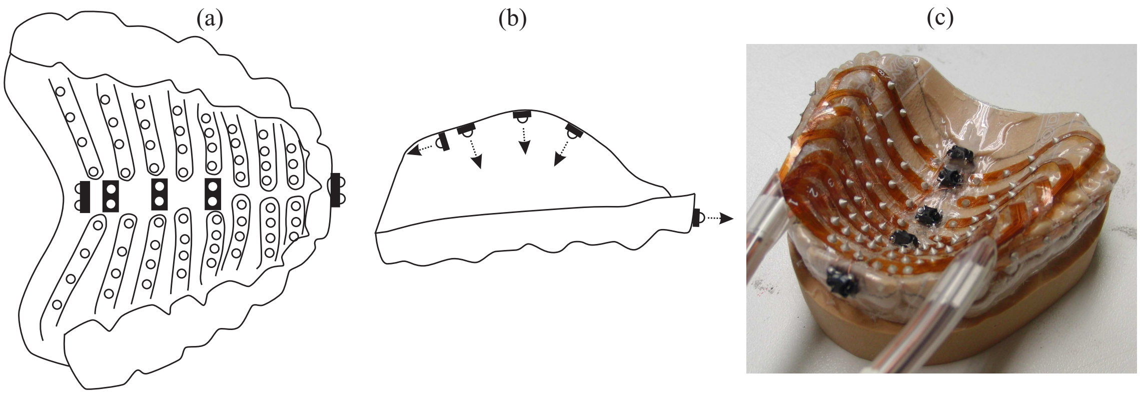

Figure 3: *(a) Bottom view, (b) side view, and (c) photography of the pseudopalate. The photography shows the pseudopalate before the the second sheet of acrylic was superimposed.*

4. Measuring system

The measuring system for the sensor data is shown in Fig. 4. The contact sensor data were measured using the WinEPG system by Articulate Instruments, and the distance sensor data were measured with a custom-made unit containing a microcontroller and conditioning circuits for the sensors. The microcontroller switched the sensor LEDs in sequence using general purpose I/O (GPIO) pins with connected current drivers and acquired the amplified reflected light data by analog-to-digital converters (ADCs). Both the WinEPG system and the custom-made unit acquired the data at a rate of 100 Hz and send them to a laptop via serial interfaces. A custom-made software on the laptop combined the two data streams to a quasi-synchronous sequence of states that could be recorded, displayed, and analyzed using a graphical user interface. For the distance sensors directed towards the tongue surface, the sensor output values were converted to distance values by means of a linear approximation of the logarithmic distance curve measured for turkey muscle tissue as shown by the dashed line in Fig. 2. However, the offset of the line was subsequently adjusted individually for each sensor. Therefore, we used the sensor output values when the tongue was slightly touching the sensors, corresponding to a distance of 0 mm.

5. Initial performance tests

For initial tests of the device, we recorded several utterances containing the sustained vowels /e,ə,ɛ,a,o,u/ and the consonants /n,m,ŋ,f,s,ʃ/ in the symmetric context of the vowels /u,i,a/. Figure 5 shows the tongue-palate contact patterns and the tongue profiles obtaines for selected vowels and consonants taken from the middle of the corresponding phones.

With regard to the vowels, there is virtually no information about the articulation of /ɛ/ and /a/ in the EPG patterns, but the tongue profiles properly reflect the different vertical tongue positions of the shown vowels. However, the current arrangement of the sensors only captures a very small part of the tongue contour. With regard to the recorded consonants, the measured EPG patterns were as expected. For example, the different places of articulation for /s/ and /ʃ/ can be clearly recognized in Fig. 5. For /ŋ/, the tongue had almost no contact with the EPG sensors, and for /n/, the sensors properly registered the full closure in the alveolar region. The measured tongue contours for the consonants look not fully convincing, especially the sharp bends in the contours of /n/ and /s/. This indicates that the calibration of the sensors must be improved, i.e. the function that relates the sensor output to distance.

The rear-directed distance sensor at the posterior end of the palate turned out to be not effective for the detection of differences in the velum position. The sensor output value was equally low for phones with a lowered velum like /n, m/ and for phones with a raised velum like /s, ʃ/. Apparently, very little light was reflected from this surface in both cases, probably because of the close-to-90 angle between the light beam and the surface normal. On the other hand, the forward-directed lip sensor in front of the upper incisors was very effective for the discrimination of unrounded vowels, rounded vowels, and bilabial closures (Fig. 5 bottom). For bilabial closures, even the context vowel can be distinguished. The discrimination of unrounded vowels and labio-dental fricatives seems not to be effective by this sensor alone.

In general, the distance sensors were not perceived as very disturbing by the user, although they protruded about 2 mm from the pseudopalate surface into the mouth. In our tests, they did not prevent the tongue from touching the EPG contact sensors in the direct neighbourhood. Deliberate coating of the sensors with saliva changed the sensor output rather little (in the range of 1–2 mm).

6. Conclusions

Our preliminary results suggest that the combination of EPG and distance sensing is very effective for the analysis of articulation. Besides immediate strengths of the prototype like the surprising performance of the lip sensor and the relative insensibility of the distance sensors against saliva, we identified many points for improvement: (i) The rear-directed sensor to detect the velum position was not effective and can be discarded in future prototypes, (ii) four distance sensors directed towards the tongue should be distributed along the entire midline (from the 1st to the 8th contact row) for more detailed tongue contours, (iii) a reliable calibration system must be designed, (iv) other types of distance sensors should be compared with the current type. Furthermore, the technique must be tested more extensively and the accuracy of measured tongue contours should be assessed by comparison with established methods like EMA. Very important from a practical point of view is that the manufacture of this type of pseudopalate is only about 3 hours more than for a conventional EPG palate, and can be further reduced

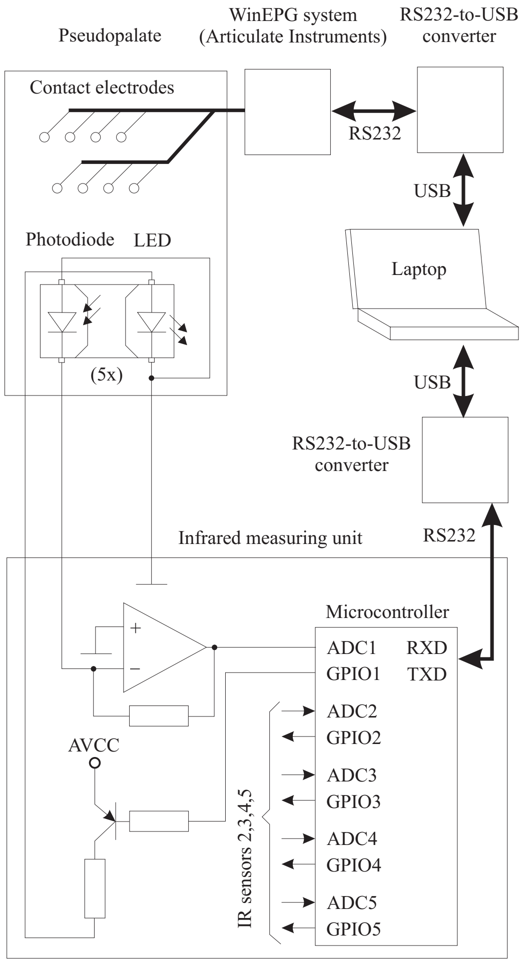

Figure 4: *Measuring system.*

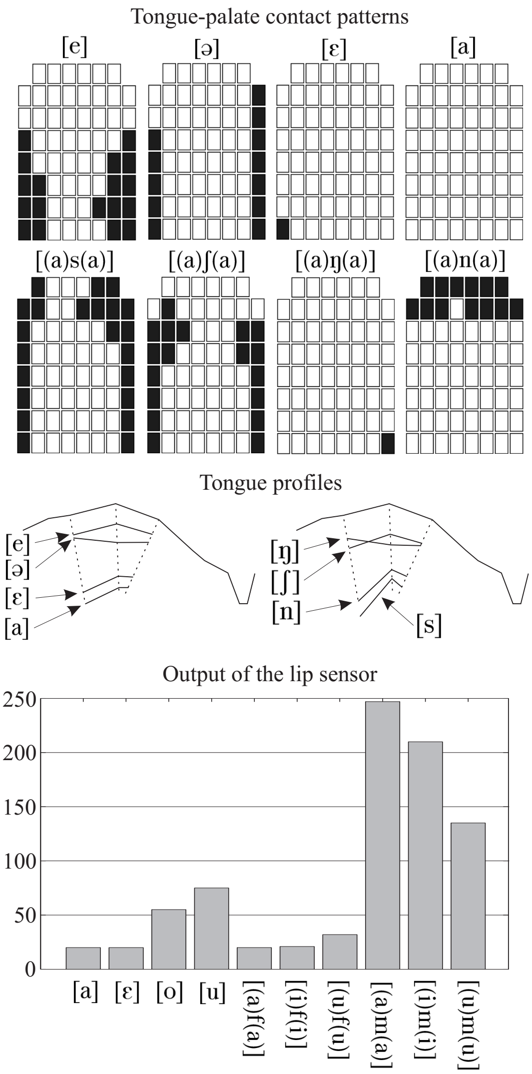

Figure 5: *Initial test results. The consonants, for which tongue profiles are shown, were spoken in the context of the vowel /a/.*

by appropriate prefabrication techniques.

7. Acknowledgements

We wish to thank Alexander Füglein for making the plaster models of the mandible, Alex Röll and Robert Nopper for their support with the manufacture of the prototype palate, Phil Hoole for the flexible circuit stips with the EPG contacts, and Doris Mücke and Martine Grice for giving us access to the WinEPG system at the Institute of Linguistics at the University of Cologne. Alan Wrench kindly provided us with detailed information about EPG palate types, palate making, and the WinEPG system, which we greatly appreciate.

8. References

[1] W. Hardcastle, W. Jones, C. Knight, A. Trudgeon, and G. Calder, "New developments in electropalatography: A state-of-the-art report," *Clinical Linguistics and Phonetics*, vol. 3, no. 1, pp. 1–38, 1989.

[2] M. M. Earnest 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*, vol. 30, pp. 2–25, 2003.

[3] C.-K. Chuang and W. S. Wang, "Use of optical distance sensing to track tongue motion," *Journal of Speech and Hearing Research*, vol. 21, pp. 482–496, 1978.

[4] S. G. Fletcher, M. J. McCutcheon, S. C. Smith, and W. H. Smith, "Glossometric measurements in vowel production and modification," *Clinical Linguistics and Phonetics*, vol. 3, no. 4, pp. 359–375, 1989.

[5] A. A. Wrench, 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)*, Philadelphia, PA, USA, 1996, pp. 1589–1592.

[6] A. A. Wrench, 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.

[7] A. A. Wrench, "Advances in EPG palate design," *Advances in Speech-Language Pathology*, vol. 9, no. 1, pp. 3–12, 2007.