EXTRACTING TONGUES FROM MOVING HEADS

Andreas Zierdt*, Phillip Hoole*, Masaaki Honda†, Tokihiko Kaburagi†, Hans G. Tillmann*
**Ludwig-Maximilian Universität München, Munich*
† NTT Basic Research Labs., Atsugi, Japan

ABSTRACT

Although existing electromagnetic mid-sagittal articulography Systems have provided invaluable information on speech kinematics, they are not suited to monitor substantial lateral deviations of the tongue from the midline without error. Due to their need for a fixed head-position they also restrict the subjects' freedom of movement and thus potentially compromise the naturalness of their speech. Accordingly we started the development of a new electromagnetic articulography system at the IPSK, which is able to acquire data within a spherical area without further restrictions. The new 3-D-EMA system is based on a spherical placement of 6 transmitter coils but operates with the same receiver coils as the Carstens AG100 system. We significantly reduced the number of analog components and built a digital-featured device. During the last couple of months we did several 3-D-Measurements to estimate the measurement accuracy of the new device. We obtained a spatial resolution better than 1 mm and a rotation detection with approx. 1 degree accuracy. We were also able to verify that the magnetic field strength of the new device is about 2 µT at the center of the measurement area which is comparable to former model calculations of magnetic field strength. We present first 3-D measurement data on a human subject collected with a prototype of the new 3-D-EMA system. To compensate movements of the subjects head during the measurement and to transform the measured data into a skull-fixed coordinate system, reference coils must be placed on the subjects head and an extra algorithm has to subtract head-movements from the movements of the sensors. Although we had to deal with some technical problems caused by preliminary parts of our prototype, the data collected to date shows that the new 3-D system is capable to aquire multi-dimensional data of speech movements. Especially the limited measurement frequency of the prototype (25 Hz) prevented us from collecting more articulographic significant data, since with only a few measured points per utterance it is difficult to analyze trajectories of speech movement. Still, even this prototype shows promising prospects and a new set of hardware components is currently being assembled. In the near future we will focus on the calibration procedure and – with increased number of channels – on the compensation of head movements during the measurements with sufficient accuracy to obtain more phonetically relevant data.

1. INTRODUCTION

Existing electromagnetic mid-sagittal articulography (EMA) Systems have provided invaluable information on speech kinematics, especially for the tongue.

Two main drawbacks remain, however. Firstly, phonetically relevant tongue movements are of course inherently three dimensional. Secondly, great care must be taken to ensure that the measurement plane of the apparatus and the mid-sagittal plane of the subject remain in alignment. This places severe restrictions on subjects' freedom of movement and thus potentially compromises the naturalness of their speech.

A related problem is that substantial lateral deviations of the tongue from the midline cannot be captured without error.

Accordingly we started the development of a new electromagnetic articulography system at the IPSK, which is able to acquire data within a spherical area without further restrictions and that derives not only three-dimensional spatial coordinates, but also the orientation of the sensors. This system is about to leave the prototype-state and become applicable for measurements on human subjects.

To obtain data on human subjects one has to estimate the movement of the subject's head to transform the measured data into a skull-fixed coordinate system. This has to be the first task when working with real-world data. To perform this, reference coils must be placed on the subjects head and an extra algorithm has to subtract head-movements from the movements of the sensors fixed to the tongue. In this way data can be obtained that is comparable with classic 2-D EMMA data, thus providing a first approach to assessing performance under realistic conditions, but avoiding many of the artificial restrictions on subject's freedom of movement in the 2D system.

2. THE 3-D EMA-SYSTEM

2.1. General description

2.1.1. 3-D-EMA-Basics. Recent EMA-Systems operate with 2 or 3 EM-field generating coils and small receiver coils which are glued e.g. on the tongue to perform fleshpoint measurement of articulatory movements. Since the field generated by a coil has dipole character, the received signal varies not only with distance between transmitter and receiver coil, but also with the angle between transmitter and receiver axis.

While the first effect yields the calculation of the distance through the "one over cubed distance"-formula, the second is most unwanted and results in the restriction to a planar measurement area for these systems.

When we started the development of the 3-D-EMA system, our goal was to use the same principles but to achieve full spatial measurement only by increasing the number of transmitter coils. In order to acquire data within a cubic or spherical area without further restrictions based on a dipole sensor, one has not only to determine the three X, Y and Z co-ordinates but also the 2 angles that describe the alignment of the dipole.

From that point of view it is absolutely legitimate to call it even a 'five-dimensional' system.

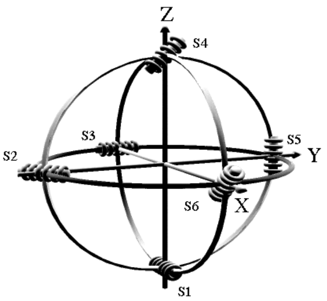

Fig. 1: Spherical placement of 6 transmitter coils.

We based the system on a spherical placement of six transmitter coils (as shown in Fig. 1) with different alignments to ensure that we always get enough information to calculate both sensor position and orientation. In mathematical terms, we have a set of six equations with five unknowns, which is solved numerically through a combined Newton/Householder-Algorithm.

2.1.2. Technical overview. Although it is obvious that the Carstens AG100 articulograph is the technical predecessor of our new system, a lot of components have changed. Because of the massive evolution in the field of Personal-Computers and digital signal-processing, we were able to develop a new digital-featured device. So we reduced the proportion of analog circuits to the absolute minimum, aiming to obtain more flexibility in the area of signal processing. We do use the AG100's amplifier circuits and the transmitter coils and resonance circuits (with minor changes) due to their well known electrical properties. We also use the same receiver coils as with the AG100.

The 3-D-EMA comes with a new 'helmet' which is more cabin like and allows free head-movements within the measurement area.(Fig. 2)

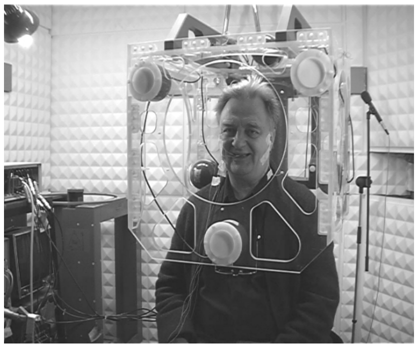

Fig. 2: The new 'helmet' of the 3-D-EMA system.

2.2. Recent data

During the last couple of months we did several 3-D-measurements with the Carstens Autokal step-motor calibration device to estimate the measurement accuracy of the new device. The very first of these results were published in Zierdt et al [3]. We found that the recent setting archives a spatial resolution better than 1 mm (0.7 mm when perfectly calibrated) and a rotation detection with approx. 1 degree accuracy.

We were also able to verify that the magnetic field strength of the new device is about 2 µT at the center of the measurement area. The final revision of the device probably has to be fine-tuned to make sure that the magnetic exposure stays below the U.S. legal limit of 205 µT even for the subjects chest (Fig. 2). We don't expect any problems in that area, because former model calculations of magnetic field strength operated with similar values, so there will be no need for significant changes of the transmitter currents.

2.2.1. Prototype specific limitations. The 3-D system we used to perform the recent measurements has still to be considered as a prototype, because of its preliminary parts. This results in some restrictions for the measurements we present here.

Since we aquired the raw data with a commercial A/D-Board with only 12-bit converter and an unfavorable gain setting, we had to accept a reduced SNR which limits spatial resolution to approx. 1 mm.[1]

The preliminary transmitter hardware caused signal drop outs every 15 to 30 minutes.

Because of a damaged input circuit, only 4 channels were available for measurements.

A missing link between transmitter and receiver clocks results in an unknown phase lag of the received signal. This occasionally caused ambiguous input data for the position-calculating algorithm, which had to be corrected offline in a time-consuming process.

The Receiver-PC we used was to slow for continuous data acquisition over a long time period, so we had to limit the acquisition time to 2.5 s per sweep.

---

[1] This was no problem with former measurements with the Carstens Autokal device, where the used measurement area was much smaller, but now we had to decrease the input gain to avoid saturated signals.

Due to the long time windows we used for the demodulation to compensate at least partly noisy data, the measuring frequency was limited to 25 Hz.

System calibration has still to be improved. It is a challenging task to select a set of calibration measurements of manageable size that will count for all possible sensor points and orientations.

3. MEASUREMENT DATA ON HUMAN SUBJECT

3.1. The setup

Two sensors were applied to the subject's head to monitor head movements (right and left side, between temple and ear). The axes of these coils were crossed to form an 'X' shape in the sagittal plane.

The other two sensors were placed on the tongue and oriented perpendicular to each other. (Fig. 3)

Simple head and tongue movements were recorded at 24.4 Hz.

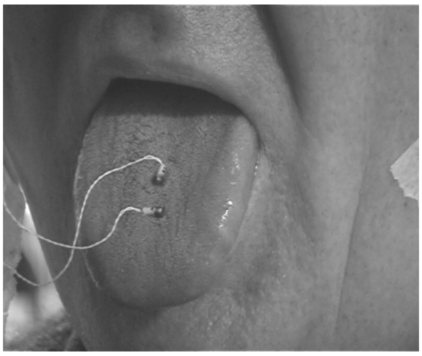

Fig. 3: Placement of the 2 sensors on the tongue.

3.2. Results

3.2.1. Sample data for nodding-movement Fig. 4 shows the movement of a sensor fixed on the subjects tongue during a nodding movement of the head. The calculated trajectory is plausible when compared to the VCR-Recording of the investigation. Fig. 5 shows the same data in an axial view. Ideally, one would expect to see just a line on the X-axis.

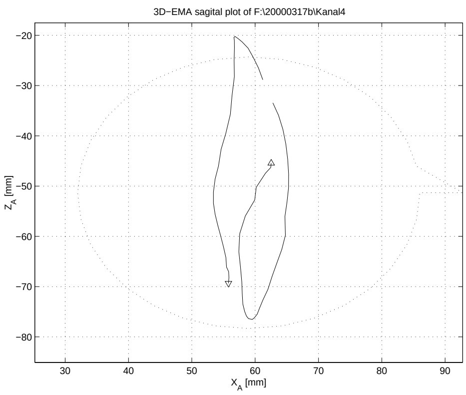

Fig. 4: Nodding head-movement, sagittal view. The dotted skull-shape illustrates only the orientation of the view, it's neither proportional nor aligned to the actual subjects head.

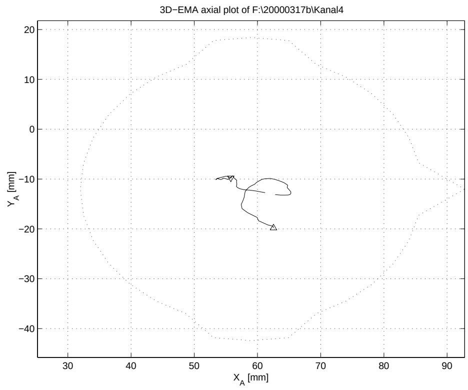

Fig. 5: Nodding head-movement, axial view.

Fig. 6 shows changes in the sensor orientation during the nodding. Nodding can roughly be described as rotation of the head around a virtual axis between the ears. Because this sensor is perpendicular to the rotation axis it shows sinusoidal changes of the orientation angles.

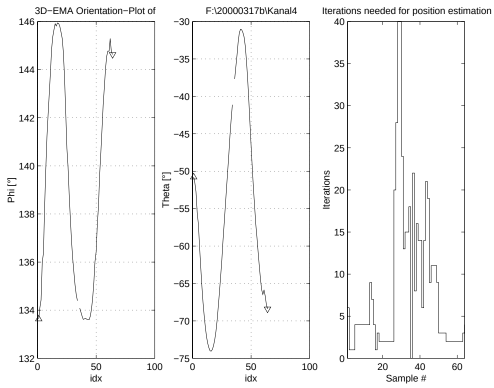

Fig. 6: Nodding head-movement, orientation of a sensor parallel to the tongue-body. Right panel shows iteration depth.

3.2.2. Sample data for speech movement. Fig. 7 gives an example of real-world data with compensated (small) head-movements. Arrows indicate the sensor orientation for the actual measurement point The figure shows the trajectory of a sensor on the tongue during the utterance 'tata'.

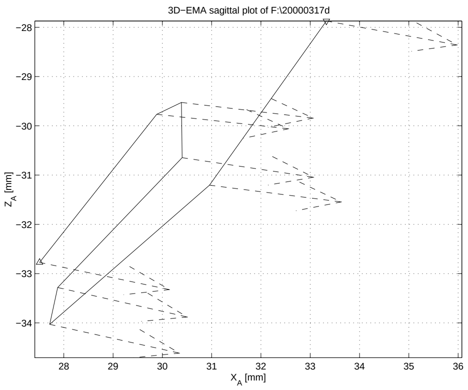

Fig. 7: Data of speech movement (utterance 'tata'). The arrows show the sensor orientation for the actual measurement point.

3.2.3. Combination of translation and rotation. Fig. 8 shows a nodding head-movement along with the sensor orientation for each measurement point. As expected the figure shows that the sensor orientation points more 'down' when the chin is moved towards the chest. Both co-ordinates are clearly related.

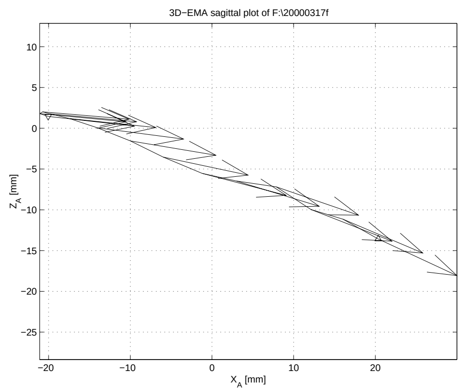

Fig. 8: Changes of sensor orientation during a nodding-movement of the head.

3.3. Conclusion

The data collected to date shows that the new 3-D system is potentially capable to aquire multi-dimensional data of speech movements, along with the ability to compensate head-movements during the measurement. It also shows that a fair amount of effort still has to be made – especially in the field of hardware – to achieve an acceptable degree of precision. The limited measurement frequency of the prototype emerged as the main drawback when evaluating the data, since with only a few measured points per utterance it is difficult to analyze trajectories of speech movements.

Still, even the prototype shows promising prospects in the field of multidimensional articulographic data collecting.

4. OUTLOOK

As soon, as the new release of hardware for the 3-D-System is available[2] we will repeat measurements on human subjects. We will focus on the calibration procedure and if necessary 'tune-up' our mathematical model to cope with the problem of the voltage-to-distance function – a situation that is familiar from the old 2-D system, as described by Hoole [1] and Kaburagi & Honda [2]

With increased number of channels and without the restricted measurement frequency and spatial resolution of the prototype, we will be able to compensate head movements during the measurements with sufficient accuracy to get more phonetically relevant data.

We also plan to perform additional MRI-investigations to further evaluate reliability and accuracy of the new device.

ACKNOWLEDGMENTS

Our work is supported by DFG Project TI 69/30 and NTT Basic Research Labs.

REFERENCES

[1] Hoole, P. (1996). "Issues in the acquisition, processing, reduction and parameterization of articulographic data", FIPKM, 34, 158-173.

[2] Kaburagi, T. & Honda, M. (1997) „Calibration methods of voltage-to-distance function for an electromagnetic articulometer (EMA) system". JASA, 101 (4), 2391-4.

[3] Zierdt, A., Hoole, P. and Tillmann, H.-G. (1999). „Development of a system for three-dimensional fleshpoint measurement of speech movements". Proceedings ICPhS99 (73)

[2] Probably at the time of the meeting.