Measurement Accuracy in 3D Electromagnetic Articulography (Carstens AG500)

Christian Kroos

MARCS Auditory Laboratories, University of Western Sydney, Australia
E-mail: `c.kroos@uws.edu.au`

Abstract

In this study the measurement accuracy of the 3D EMA (Electromagnetic Articulograhy) system Carstens AG500 was evaluated using an optical motion capture system (Vicon). A custom-built container fixed the positions of the EMA sensors and the optical markers and a series of movements of the container within the EMA measurement field was recorded. The results show that (a) the very small error of static position measurements does not reflect the behaviour of the system in dynamic measurements and (b) that in dynamic measurements the magnitude of the error varies substantially across the three spatial dimensions.

1 Introduction

Since the beginning of articulatory phonetics as an experimental science, measurement problems have been its unwanted companions. Most of the behaviour of the human articulatory organs is not directly visually observable, speech is inherently dynamic, and measuring articulatory movements strictly requires that measurement procedures do not interfere with their execution. Hence the choice of appropriate methods is limited and almost always accuracy of capturing one aspect is achieved at the expense of neglecting other aspects. This applies especially to measurements of the moving tongue. Cineradiography, Electropalatography (EPG), structural Magnetic Resonance Imaging (MRI) and ultrasound all provide data about aspects of tongue behaviour in speech, but only the recently developed three-dimensional Electromagnetic Articulography (EMA) measures tongue movements in 3D and real time returning data from tongue surface flesh points. Electromagnetic Articulography started almost forty years ago with very limited 2D systems [1, 6] that were during the next decades substantially improved and refined [5, 2, 3] until in the late 1990s a 3D system was developed [7]. Currently two commercial systems are available, the AG500 (Carstens Medizinelektronik) and the Aurora system (Northern Digital). At the time of the writing, the latter - originally designed for guiding instruments in surgery - does not deliver sufficient temporal resolution with the number of sensors usually required in speech production experiments.

In the AG500 six transmitter coils fixed to a cube-shaped acrylic glass structure produce an alternating electromagnetic field, each coil with a characteristic frequency. A very weak current oscillating with the same frequency is induced in a set of small sensor coils brought into the field with the voltage of the induced current depending on the distance from the transmitter coil and the orientation of the sensor. Using demodulation of the compound signal in the sensor the contribution of each transmitter coil can be identified and the spatial position of the sensor calculated. This calculation, however, is intricate. The computational problem cannot be linearised in the 3D case and iterative non-linear optimisation methods have to be used which can fail to converge on the globally best solution under certain circumstances. As a consequence, certain combinations of location and orientation of the sensors might produce larger errors and should be avoided in speech production experiments. Thus, knowing what constitutes these conditions becomes crucial in 3D EMA. Until now, however, there have been very few evaluations of the measurement accuracy of the AG500 (but see [4]), and to our knowledge no study has examined accuracy in dynamic situations with the sensors moving in a similar manner as in speech production experiments with human participants.

2 Method

In this study we used an optical motion capture system (Vicon, Oxford Metrics) with passive mark-

Figure 1: *Inside of the SMC showing the three EMA calibration cartridges*

Figure 2: *The SMC with the OPT markers at the end of the 3D cross structure extension*

ers (henceforth abbreviated OPT) and conceived and built an EMA/OPT Simultaneous Measurement Container (SMC) to obtain reference values to which the EMA measurements could be compared. The SMC consisted of a rigid plastic container to hold the EMA calibration cartridges with the sensors inserted and an extending splint to attach the OPT markers. Three holes were drilled into the base and walls of the container to secure all three EMA cartridges firmly in three different main orientations in the interior of the SMC using the plastic screws of the AG500 calibration unit (see Figure 1). A non-flexible plastic splint was glued to the underside of the SMC extending roughly five centimetres beyond the front side of the SMC and a piece of plastic shaped in the form of a three-dimensional cross was put over the extended part and glued to the splint. The cross-like structure allowed anchoring the heads of small plastic screws in narrow grooves along its beams which were additionally fixed using superglue. Spherical 9.5 mm OPT markers were screwed to their protruding end (see Figure 2). In this way all 12 EMA sensors and 8 OPT markers were part of the same single rigid object, the SMC, and every movement of the SMC affected all sensors and markers in a predictable way.

Before the experiment the AG500 was painted black using acrylic matte black paint to avoid reflections that would disturb the optical motion tracking. Four Vicon MX40 cameras were placed around the EMA cube (see Figure 3).

Both systems were calibrated and the size of the residuals indicated good calibrations. The RMS values of the individual sensors of the EMA system were well below the recommended maximum value of 20 with eleven of them between 8.6 and 11.2 and only one sensor showing a slightly higher value (14.5). At the beginning of the experiment the SMC was suspended within the EMA cube using conventional cord and five static trials were recorded. Then the SMC was cut loose and moved for the rest of the experiment by the experimenter, with his hand extended into the cube. Ten trials each were recorded with movements consisting predominantly of translational movements in each spatial dimension separately. Another ten trials each were recorded with movements consisting predominantly of rotational movements around the three axes separately. An-

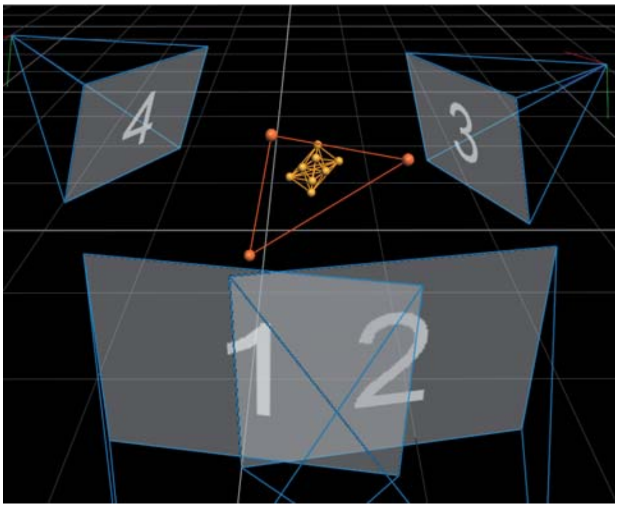

Figure 3: *Schematic representation of the Vicon setup showing the positions of the four cameras, the cube markers (red), and the SMC markers (yellow)*

Table 1: *RMS error (in mm) of all inter-marker and inter-sensor distance deviations for the different movement types. For the EMA system values for the sensors within single cartridges are also reported (last six columns). MX, MY, and MZ indicate the primary orientation of the cartridges along the X, Y, and Z axis, respectively. CP and TA refer to coordinate calculation from the raw voltage data with either Calcpos (CP) or TAPAD (TA). Distance deviations larger than 10mm were treated as missing values and excluded.*

	OPT	EMA							
		All		MX		MY		MZ	
		CP	TA	CP	TA	CP	TA	CP	TA
Static	**0.06**	**0.06**	**0.37**	**0.06**	**0.35**	**0.06**	**0.28**	**0.05**	**7.81**
Translation along X axis	0.30	0.84	1.55	1.05	1.11	0.60	0.72	0.32	5.10
Translation along Y axis	0.32	1.11	1.31	0.65	0.70	0.99	1.00	0.62	2.68
Translation along Z axis	0.37	1.06	1.81	1.44	1.50	0.83	1.02	0.33	4.07
Rotation around X axis	0.45	1.45	1.56	1.61	1.50	1.10	1.11	0.90	5.03
Rotation around Y axis	0.34	1.44	1.60	1.77	1.84	0.83	0.79	0.52	1.10
Rotation around Z axis	0.43	1.66	2.08	1.74	1.81	0.95	0.96	0.67	3.98
Unconstrained	**0.37**	**1.54**	**1.72**	**1.61**	**1.56**	**0.81**	**0.90**	**0.76**	**2.09**
Average	**0.33**	**1.15**	**1.50**	**1.24**	**1.30**	**0.77**	**0.84**	**0.52**	**3.98**

other 20 trials were recorded with no restriction on the movement type with a relatively small movement range comparable to the range usually found in speech articulator movements. Temporal synchronisation was achieved by recording the ’sweep’ (on/off) signal generated by the AG500 using Vicon’s synchronised analog signal recording capabilities. All trials consisted of 2500 samples recorded with both systems with a sample rate of 200Hz. Finally five trials were recorded without the SMC but with three EMA sensors placed exactly in the hollow centre of three OPT markers and then suspended in the cube to determine the global offset between the two systems. Cube movements were tracked during the whole experiment with an additional three 14mm OPT markers that were attached to the EMA cube exploiting existing small holes in the cube and similar plastic screws as used in the SMC. The EMA position coordinates and angles were computed with the routine provided by the manufacturer (Calcpos) and alternatively with the TAPAD Matlab toolbox developed by Andreas Zierdt at the Phonetics Department of Munich University.

All movement data were low-pass filtered and downsampled to 50Hz. The distances between the sensors/markers within both systems (which would be constant in a noise-free measurement system) were calculated as a first evaluation measure. The spatial coordinates and orientation angles (not reported here) of the EMA sensors were estimated based on the OPT data and the average sensor location in the static trials (using Calcpos). The mean squared error (RMS) between the estimated and the measured EMA data was computed as the main accuracy measure.

3 Results and Discussion

The OPT system returns missing values if it cannot determine the location of a marker (the criteria for the exclusion are not reported by Vicon). while the EMA system always returns a value even if it is out of range. Since these values would distort a realistic assessment of the measurement accuracy it was decided to exclude distances sample-wise that exceeded a threshold deviation of 10mm - a value which was considered big enough to be detected in speech production experiments as a system error and thus could be declared missing. Table 1 shows the deviation of the within-system inter-marker/inter-sensor distances from their static means. Except for the static trials the RMS error is approximately three times smaller in the OPT system as in the EMA system, justifying the use of the former to evaluate the latter.

Table 2 gives the RMS error for the deviation of

Table 2: *RMS error (in mm) characterising the deviation of the position coordinates measured with EMA from the ones predicted using the OPT system. The RMS error in terms of the Euclidean distance is given in the first two value columns, the RMS error for the deviations in the three spatial dimensions separately in the remaining six value columns. CP and TA refer to coordinate calculation with either Calcpos (CP) or TAPAD (TA). Samples of single sensors producing an RMS error larger than 10mm were treated as missing values.*

	Euclidean		X		Y		Z	
	CP	TA	CP	TA	CP	TA	CP	TA
Static	**0.16**	**0.74**	**0.04**	**0.30**	**0.06**	**0.61**	**0.15**	**0.30**
Translation along X axis	4.54	4.58	4.15	4.12	0.66	0.82	1.71	1.83
Translation along Y axis	4.22	4.19	3.24	3.10	2.43	2.51	1.18	1.27
Translation along Z axis	5.05	5.17	3.86	3.83	0.95	1.14	3.11	3.27
Rotation around X axis	5.02	5.14	4.16	4.15	1.54	1.48	2.36	2.64
Rotation around Y axis	5.17	5.15	4.32	4.18	1.07	1.14	2.64	2.77
Rotation around Z axis	5.33	5.45	4.74	4.66	1.91	2.16	1.54	1.82
Unconstrained	**4.20**	**4.38**	**3.24**	**3.21**	**1.82**	**2.06**	**1.95**	**2.15**
Average	**4.21**	**4.35**	**3.47**	**3.44**	**1.31**	**1.49**	**1.83**	**2.01**

the position coordinates measured with EMA from the ones predicted with the OPT system. Note that this implies that any inaccuracy of the OPT system might be responsible for a portion of the observed error. Again, samples of single EMA sensors that produced an RMS error higher than 10mm were treated as missing values. A large difference between the static trials and all movement trials can be seen with the error in the X axis measurements being the primary factor responsible for the increase in error magnitude. The summary table does not reflect an observed substantial underlying variability of the RMS error dependent on location and orientation of the sensors. Note that the RMS error penalises for larger deviations and thus passages and/or single sensors with high deviations had an pronounced impact on the reported summary values. However, space limitations prevent us from presenting a detailed analysis here.

4 Conclusion

The results of the evaluation of the EMA system Carstens AG500 presented in this study show that careful consideration of the sensor placement and orientation is required in order to keep the measurement error within an acceptable range. In future work we will map the error in detail and with respect to the most frequently used locations and orientations of EMA sensors in speech production experiments.

Supported by NHMRC/ARC grant TS0669874 and the AHAA UWS RIF grant.

References

[1] T. J. Hixon. An electromagnetic method for transducing jaw movements during speech. *The Journal of the Acoustical Society of America*, 49(2B):603–606, 1971.

[2] P. Hoole. Issues in the acquisition, processing, reduction and parameterization of articulographic data. FIPKM, 34, 1996.

[3] T. Kaburagi and M. Honda. Calibration methods of voltage-to-distance function for an electromagnetic articulometer (EMA) system. *The Journal of the Acoustical Society of America*, 101(4):2391–2394, 1997.

[4] T. Kaburagi, K. Wakamiya, and M. Honda. Three-dimensional electromagnetic articulography: A measurement principle. *Journal of the Acoustical Society of America*, 118(1):428–443, 2005.

[5] J. S. Perkell and D. Oka. Use of an alternating magnetic field device to track midsagittal plane movements of multiple points inside the vocal tract. *The Journal of the Acoustical Society of America*, 67(S1):S92–S92, 1980.

[6] G. van der Giet. Computer-controlled method for measuring articulatory activities. *The Journal of the Acoustical Society of America*, 61(4):1072–1076, 1977.

[7] A. Zierdt, P. Hoole, H. M., T. Kaburagi, and H.-G. Tillmann. Extracting tongues from moving heads. In *5th Speech Production Seminar*, pages 313–316, Kloster Seeon, Bavaria, Germany, 2000.