# Accuracy Assessment for AG500, Electromagnetic Articulograph

**Yana Yunusova**
University of Toronto, Canada

**Jordan R. Green** and **Antje Mefferd**
University of Nebraska–Lincoln

## Abstract

**Purpose—**The goal of this article was to evaluate the accuracy and reliability of the AG500 (Carstens Medizinelectronik, Lenglern, Germany), an electromagnetic device developed recently to register articulatory movements in three dimensions. This technology seems to have unprecedented capabilities to provide rich information about time-varying positions of articulators. However, strengths and weaknesses of the system need to be better understood before the device is used for speech research.

**Method—**Evaluations of the sensor positions over time were obtained during (a) movements of the calibration device, (b) manual movements of sensors in a cartridge within the recording field of the cube, and (c) various speech tasks.

**Results—**Results showed a median error to be under 0.5 mm across different types of recordings. The maximum error often ranged between 1 and 2 mm. The magnitude of error depended somewhat on the task but largely on the location of the sensors within the recording region of the cube.

**Conclusion—**The performance of the system was judged as adequate for speech movement acquisition, provided that specific steps are taken for minimizing error during recording and for validating the quality of recorded data.

### Keywords

electromagnetic articulography; AG500; speech kinematics

Electromagnetic articulography (EMA) is rapidly becoming the predominant technology used to study movements of the tongue during speech and swallowing (Chen, Murdoch, Goozee, & Scott, 2007; Hertrich & Ackermann, 2000; Steele & Van Lieshout, 2004). The AG500 (Carstens Medizinelectronik, Lenglern, Germany) is currently the most developed three-dimensional (3D)-EMA system. This device is superior to its two-dimensional (2D) predecessors (AG100, AG200) because it does not require a participant to wear a heavy, restraining head mount and it provides motion tracking in five degrees of freedom (i.e., three Cartesian and two angular coordinates). Additionally, it is not adversely affected by midline shifts of the tongue or rotational misalignments of the sensors (Hoole, Zierdt, & Geng, 2003). Therefore, in comparison to the 2D-EMA, the 3D-EMA is expected to produce smaller measurement errors across a larger range of sensor positions and orientations (Kaburagi, Wakamiya, & Honda, 2005), providing an unprecedented level of access to the most complex lingual and labial articulatory behaviors.

*Corresponding author* – Yana Yunusova, Department of Speech and Language Pathology, Rehabilitation Sciences Building, University of Toronto, 160-500 University Avenue,Toronto, Ontario M5G 1V7, Canada; yana.yunusova@utoronto.ca.

The 3D-EMA has only recently been released for commercial use, and its accuracy and reliability are unreported. Reports of positional errors at the experimental stage of AG500 development were approximately 1 mm (0.7 mm when the system was perfectly calibrated), and rotational error was about 1 degree (Hoole et al., 2003; Zierdt, Hoole, Honda, Kaburagi & Tillmann, 2000). More detailed information is available in the literature for its 2D predecessors —the AG100 and AG200 (Hoole, 1996; Tuller, Shao, & Kelso, 1990), and the EMMA at MIT (Perkell et al., 1992). The spatial resolution of the 2D systems has been reported to be approximately 0.5 mm, which has been judged adequate to capture time-varying positions of articulators in order to monitor acoustically relevant changes in vocal tract geometry (Perkell et al., 1992). However, midline shifts combined with rotational misalignments of the receiver coils relative to the axes of the transmitters significantly increase positional errors in the 2D systems and are a particular challenge because the occurrence of midline shifts is difficult to identify during data collection (Honda & Kaburagi, 1993; Perkell et al., 1992).

In summary, the AG500 has potential to provide unprecedented access to tongue movement data during speech and swallowing. The accuracy and reliability of the device, however, have not been confirmed by independent laboratories. Therefore, the goal of this report is to document spatial resolution of the AG500 and to define the boundaries of measurement error for future reference in analyses of speech production data. The following distinct but complementary aspects of movement tracking were investigated: (a) the reliability of sensor calibrations and positional tracking over time, (b) the absolute and relative spatial error during optimal recording conditions and during speech, (c) the relationship between sensor calibration values and the magnitude of positional errors, and (d) the uniformity of tracking accuracy within the recording volume.

## Method

### Basic Principles of AG500 Operation

The principles of the 3D-EMA have been investigated thoroughly and described elsewhere (Zierdt, 1993; Zierdt, Hoole, & Tillmann, 1999; Zierdt et al., 2000). Briefly, the AG500 is a system of six transmitter coils (see Figure 1) arranged spherically such that each receiver coil (sensor) axis is never perpendicular to more than three transmitters at once. The transmitters are driven at different frequencies, ranging from 7.5 to 13.75 kHz. Each transmitter electromagnetically induces currents in up to 12 receiver coils (sensors). The voltage measured at the sensors varies as a function of the distance and the angle between the axis of each transmitter and as a function of each sensor. The AG500 quantizes induced voltage values (amplitudes) at 16-bit resolution. The six measured amplitudes are used to calculate the distance between each transmitter and each sensor, taking into consideration the sensor's angular coordinates. Cartesian and angular coordinates of each sensor are determined, then, by solving a set of complex mathematical equations. Next, the expected amplitudes are derived for each calculated position in the measurement field. This reverse computation is based on the known field model—a mathematical representation of the spatial pattern of the magnetic field in the measurement volume (see Zierdt, Hoole, & Tillman, 1999).

Measured and expected amplitudes, derived based on the reverse computation, are then compared, and the root-mean-square (RMS) error value representing the difference between the two sets of amplitudes for each sensor position are calculated. If the measured and expected amplitudes are identical, the RMS error is zero. The system accepts calculations with RMS errors under 62 digits, with one digit corresponding to one quantal level (at 16-bit resolution, 65,536 digits correspond to a 5-volt input). However, this RMS threshold can be adjusted by the experimenter. If the RMS value is above the threshold, positional calculations are repeated until a better match between the measured and expected values are found and the RMS threshold is reached. Sometimes, more than one positional solution is stored during these

calculations, and the experimenter is allowed to choose a preferred solution based on a graphical visualization of the computed positions. The system also provides the user with access to the RMS values for each calculated position and a statistical summary of these values across multiple positions for each sensor. This information can then be used for evaluation of the quality of recordings.

## Calibration

The system requires calibration. The calibration is performed to determine parameters of the field model and to define the relationship between voltage and distance for each transmitter coil and each sensor. During the calibration procedure, 12 or fewer sensors are attached to a rotating disk (*circal*) via cartridges (see Figure 1) and are rotated by a motor incrementally in 8,000 steps over a full circle (360°). The disk, which is oriented parallel to the axial (vertical) plane, is rotated around its center. The center of the disk is located a few millimeters above the cube's origin, which is the middle of the 3D volume circumscribed by the EMA cube. Prior to the beginning of the calibration, a position called *Logic Zero* has to be identified. Logic Zero specifies the position at which Sensor 1 crosses the positive *x* axis. The calibration software then runs automatically for approximately 45 min (less time is required in the latest version of the program). Following calibration, the system software provides a number of output summary parameters that are used to judge the quality of the calibration. These parameters include *calibration factors* (measured amplitudes), deviations of each sensor from the Logic Zero position (called *alpha-zero*), and average RMS calculated across all positions for each calibrated sensor. Additional reference values for each sensor include estimates of the circal radius (R) representing combined X and Y dimensions, circal Z (Z coordinate), and orientation angles phi ($\Phi$; tilt) and theta ($\theta$; yaw). Reference values for each parameter are provided in the *AG500 Sensor Calibration Manual*. Calibration factors should be between 2,100 and 2,400 digits, alpha-zeros should be randomly distributed around zero and be between $\pm 0.5°$, and RMS values should be fewer than 20 digits. R, Z, $\Phi$, and $\Theta$ values are assumed to be constant throughout the circal for each sensor. R should be approximately 80 mm, and the Z coordinate should be approximately 6.5 mm above the origin of the EMA cube. The sensor's $\Phi$ should be approximately 45°, whereas $\Theta$ should be less than 5°. If one or more of these parameters are outside the expected range, the manufacturer recommends sensor recalibration until acceptable values are achieved. To improve on the calibration results, it is recommended that (a) the sensors be checked for damage, (b) the sensor be realigned in the cartridge, (c) the Logic Zero be readjusted, and/or (d) the system be warmed up sufficiently (up to 3 hr) prior to the next calibration.

## General Procedures

After moving the system through various locations in the lab and monitoring changes in the sensor amplitudes using diagnostic software, the AG500 device was positioned such that environmental influences (e.g., proximity to walls, metal interference, fluorescent lighting) were judged to be minimal. As recommended by the manufacturer, the device was allowed to warm up for at least 3 hr prior to each calibration and recording. Calibrations were performed prior to each new recording (with an exception of repeatability of circal recordings). For all but one analysis (distance between sensors on the jaw), sensors were secured tightly in the cartridge (see Figure 1) to prevent slippage during testing. For recordings where sensors secured in the cartridges were moved manually within the cube, the experimenter was grounded. Only HS220s sensors were used for all of the recordings. Although they have been replaced recently by new sensors (HQ220s), HS220s are still in use in a number of laboratories. The manufacturer reports that HQ220s produce a stronger signal and better signal-to-noise ratio. Our informal observations revealed similar results in quality of positional tracking between the two types of sensors.

Movement data were collected and processed following instructions provided by the manufacturer. Movement signals were acquired at 200 Hz. The software program CalPos_2 was used for position calculations (another software program [TAPADM], developed by Andreas Zierdt, is also available for position calculation). Prior to analyses, the movement channels were low-pass filtered at 10 Hz using a zero-phase shift forward and reverse digital filter (Butterworth, 8-pole). Methodological details specific to the different analyses are provided in each subsection of the Results section.

## Results

### Sensor Calibration

Results of 12 consecutive calibration runs across all sensors collected during a period of testing of the system between February and December 2006 were combined for this summary. Usually, sensors were calibrated two to three times before the calibration results were acceptable. Only results for runs and sensors with calibration RMS, calibration factors, and alpha-zeros within the range suggested by the manufacturer are reported ($n = 130$). In summary, the median calibration RMS across sensors and multiple calibration runs was 12.78 digits, with a minimum and maximum of 8.22 and 16.52 digits, respectively. The interquartile range (IQR) of this distribution was 3.15. Approximately 75% of the calibration RMS values were under 14 digits. Based on these results, we adopted a calibration criterion of RMS = 14 digits as a cutoff value for an acceptable calibration for our system. This value is more stringent than the manufacturer's recommended value of RMS = 20. For our AG500, we now recalibrate the sensors until the value of 14 or better is reached; our system achieves this RMS level in the majority of first calibration runs. The analyses reported below were performed only on the subset of data with RMS equal to or less than 14 ($n = 103$) unless stated otherwise.

### Spatial Error Estimates for Circal Movement Recordings

**Absolute error: X, Y, Z, R, Θ—**Movement of the sensors in the circal was recorded immediately after each calibration, with position and orientation of sensors unchanged in the magazines after the calibration. X, Y, Z, R, and Θ vectors in the recorded circle were compared with their respective expected values, which were calculated during the calibration procedure and given in the calibration summary file. For this analysis, it was assumed that the sensor movement path was a perfect circle with constant dimensions for Z, R, and Θ. Predicted X and Y vectors were modeled as cosine and sine functions:

$$[X_{MODEL} = R \times \cos \alpha]\,[Y_{MODEL} = R \times \sin \alpha]$$

The sine and cosine functions were modeled in 3,600 steps over $2\varPi$ radians, which provided 10 data samples per degree. In order to make the comparison between the measured positional data and the models, the measured data were down-sampled to match the sampling frequency of the models and were aligned with the models using the angle $\alpha$ as a reference. The error was calculated as the absolute difference between the measured and estimated X and Y coordinates for the down-sampled vectors. Additionally, the values of Z, R, and Θ estimated during the calibration procedure were subtracted from the measured Z, R, and Θ vectors to determine the error in these parameters.

An example of the analysis is shown in Figure 2. On the top plots, the modeled/expected functions are plotted as thin black lines, and the measured data are represented by thick, light-gray lines. The bottom plots show error functions representing the difference between measured and expected positions. For all plots, angular positions ($\alpha$, in radians) at which the modeled and measured data were aligned and compared are plotted on the *x*-axis. The positional

(X, Y, Z) and error data are plotted on the *y*-axis. The median, IQR, and maximum of absolute values of each error function were calculated for each sensor and session. The summary statistics presented in Table 1 show the grand median, IQR, and range of errors in each dimension calculated across records. The grand median for each signal appears to be relatively small (0.22–0.39 mm). The variability expressed as IQR is approximately 0.4 mm for the unidimensional signals (i.e., X, Y, Z). The median of maximum error had a consistent range across recordings of 1.5–2.5 mm. Median errors in the sensor's Θ appear very small (0.2°–0.3°), with the maximum being under 1°.

**Relative error: Distance between sensors in circal—**To estimate 3D spatial error in contrast to the unidimensional estimates presented previously, the Euclidean distance between each adjacent sensor pair within a cartridge (total of nine pairs) was calculated for three randomly selected circal recordings. Because the results of this error analysis were nearly identical for the three separate recordings, the statistical summary from a single recording is reported here. Figure 3 shows the time history of movements of two adjacent sensors in the Z dimension with each sensor reaching the same location as the preceding sensor with a short time lag. The similarity between the signals and the short time lag between them illustrates that the tracking distortions are highly location dependent. Prior to the error analysis, the time series signals of selected sensor pairs were temporally aligned to minimize the effect of the location-related distortions on our error estimates. Sensor pairs were temporally aligned based on a time lag between the signals that was computed algorithmically using a cross-correlation approach (see Green et al., 1997;Green, Moore, Higashikawa, & Steeve, 2000). After the signals were aligned, the distance functions between adjacent sensors were calculated. The error function was calculated as an absolute mean-corrected distance between each adjacent sensor pair. Under ideal recording conditions, the distance between the sensors should remain constant throughout the circal. Our results showed that the median error calculated across sensor pairs was 0.52 mm, with IQR of 0.36 mm and a maximum error of 1.94 mm.

## Calibration RMS Predicting Spatial Error

A correlation analysis was performed to determine if RMS values obtained during calibration predicted the accuracy of positional tracking. For this analysis, Pearson product–moment correlations were calculated between calibration RMS values and values for both median and IQR errors calculated for the X, Y, Z, and R signals across sensors and sessions (see section titled *Absolute error: X, Y, Z, R, Θ*). These correlation analyses were performed on a subset of circal recordings ($n = 103$, RMS under 14) and on the complete set of data ($n = 130$, all RMS under 20). Significant but relatively weak correlations were observed between RMS and spatial error estimates of R median error ($r = .30$) for the full set of data only. Significant but also weak correlations were observed for both the complete set and the subset of data for the median error in Y ($r = .43$ and .36, respectively) and the median error in X ($r = .39$ and .41, respectively). Figure 4 shows a plot of calibration RMS versus the median error in Y for the full data set ($n = 130$) with a best-fit regression line.

## Effect of Position in the Cube on Magnitude of Errors: "Small Movement" Analysis

The manufacturer of the AG500 recommends that a "small movement" procedure be performed as an additional test of quality of sensor calibration. The analysis of data collected during circal recordings clearly demonstrated that spatial error is nonuniform within the measurement area of the cube (see Figure 3). In this analysis, we attempted to evaluate positional errors in relation to the sensor location within the measurement field of the device. The manufacturer company defines the optimal measurement field as a sphere 15 cm in radius around the cube origin. Well-calibrated Sensors 2 and 8 (with a calibration RMS of 10.8 and 9.8, respectively) were secured in a magazine next to one another. We manually moved the magazine in small excursions, sampling the entire range of the optimal measurement field of the cube. The area was sampled

in four recordings of approximately 60 s in duration. For the purposes of this analysis, the error was defined as an absolute value of the mean-corrected distance function calculated between the two sensors. The distribution of the errors across four records had a median of 0.49 mm, with the middle 50% of data having error between 0.18 and 0.83 mm. The 5% of data at the high tail of this distribution had error of 2.43 mm and above. A portion of the error vector with positional error above 0.5 mm is plotted inside the optimal measurement field on Figure 5. The range of error is shown in grayscale gradients (in mm) on the top of each plot. The left plot shows error in the coronal (Y–Z) plane, whereas plots in the middle and on the right show error in the sagittal (X–Z) and transverse (X–Y) planes, respectively. The figure suggests that unlike the data recorded in the circal, errors greater than 0.5 mm can be distributed across the measurement field except for the central region.

### Distance Between Sensors on the Jaw

In order to assess the magnitude of positional errors during speaking, two sensors were glued to the jaw: one on the buccal surface of the lower incisors at the midline and the other off the midline on the left side between the canine and first premolar. A single speaker was asked to perform the following tasks: phonate /a/ for 5 s, repeat a syllable /ba/ for 15 s, read a sentence ("Tomorrow Mia may buy you toys again") five times at a normal, comfortable speech rate (approximately 20 s), and read a paragraph normally (approximately 60 s). The Euclidian distance between the two jaw sensors was calculated for each task. Under the best circumstances, this distance should change minimally over the course of the recording. Summary statistics for absolute mean-corrected distance function computed for each task are shown in Table 2. The median error was small and varied between 0.07 and 0.22 mm between the tasks. The error IQR ranged between 0.03 and 0.30 mm, depending on the task. Maximum error was 2.00 mm for the longest task (paragraph reading).

### Repeatability of Circal Recordings

The reliability of positional tracking for a single calibration was assessed within and across days. For this analysis, all of the sensors except Sensors 9 and 12 were calibrated, and a series of circals were recorded. Five circals were consecutively recorded on Day 2 postcalibration and one on each subsequent day for 10 days in a row. The sensors were never removed from the magazines in between recordings, and all recordings started at the Logic Zero position. The Euclidean distance from the origin (the center of the cube) was calculated for each sensor. To estimate reliability of circals on multiple recordings, the Euclidean distance functions for the first circal recording were subtracted from the Euclidean distance functions collected for each subsequent recording; the highest degree of reliability was indicated by the minimal difference between each circal run.

Results showed that the measured Euclidian distances between two sensors were nearly identical across the five repetitions recorded on the same day (Median difference = 0.02 mm). Table 3 shows the accuracy of the measured Euclidean distance between the same sensor's positions recorded on different days (e.g., Day 1 and Day 2), suggesting that the errors increase gradually across recording days.

## Discussion

The purpose of this investigation was to test the accuracy and reliability of the AG500. The results showed that the performance of the 3D electromagnetic device was reasonably good but that positional errors were unacceptable in some localized regions of the cube. For the circal recordings, summaries of positional errors revealed relatively small median errors (usually under 0.5 mm) in each dimension (X, Y, Z). The error in a single angular coordinate evaluated (θ) was well under half of a degree. The errors calculated as Euclidian distances between pairs

of sensors were relatively small, as well. For circal and "small movement" recordings, median errors were under 0.5–0.6 mm, which is small considering that combining sensors may have an additive effect on the error.

However, our results also showed a large range of positional errors across all of the analyses and types of recordings, with maximum errors up to 2 mm during circal movements and some individual recordings of circal showing error up to 4–5 mm. The circal movement analyses revealed a number of regions within the recording field of the cube that resulted in elevated error. The positions with high errors were relatively predictable, with high errors occurring at similar places along the circal movement path for each sensor and across recordings.

The analysis of small movements showed that 25% of errors were above 0.83 mm. High errors were observed to occur anywhere within the recording field of the device, perhaps with the exception of the cube center. However, it is also possible that the procedure for obtaining small movement data (i.e., manual movement of the magazine holding the sensors and relatively long intervals of recording) was not ideal for this analysis. The magnitude of error during this task might have also been affected by the relatively large movement magnitudes (as compared with speech) and/or high movement speeds (recall that the magazine was moved repeatedly through different locations within the cube). It is also possible that the error was elevated because the boundaries of the measurement field were not set, and moving outside of the boundary affected the quality of recording after the cartridge with sensors returned into the measurement space. The manufacturers have recently developed a new tool (called the "accuracy checker") for estimating the spatial errors, which will estimate the magnitude of positional error as a function of location in the cube and sensor orientation. Unfortunately, the small evaluation area of this device (70-mm radius around the center of the cube) will still leave a significant region of the field untested.

At this time, there is no other technology available that is fully comparable in capabilities with the AG500. Users of this technology, however, should be aware of its limitations and should not assume that every movement registration with the device is accurate. Based on our preliminary experience with the system and the current findings, we recommend several procedures for minimizing the magnitude of errors during data collection. The first and obvious step is to ensure the best possible calibration of the sensors. Inadequate calibration will decrease the accuracy of positional tracking (Kaburagi et al., 2005; Zierdt, 2007; Zierdt et al., 1999); more robust calibration techniques are currently being developed (Zierdt, 2007). Until this is achieved, a close monitoring of the quality of calibration recordings is imperative. We found that calibration RMS is perhaps the most useful parameter to monitor the quality of calibration, assuming all other parameters (e.g., calibration factors, alpha-zero, θ, etc.) are within expected limits. Based on the data collected for HS220s sensors, we recommend accepting calibrations only if they are relatively small (14 digits for our system). The RMS value should be stringent but also relatively easy to achieve (75% of our calibrations resulted in sensor RMS of fewer than 14 digits). The analysis of the relationship between calibration RMS and spatial errors (see Figure 4) also showed that for RMS values more than 14 digits, the relationship between spatial error and calibration RMS is rather poor, which might suggest that collecting data using sensors with these calibration characteristics should be avoided. We do not know if a similar value (14 digits) will be obtained for other systems. Calibration RMS might also be different for the newly released sensors and will need to be re-estimated as data collected using those sensors become available.

Additionally, the results of our reliability testing, which consisted of comparisons of circal movements on repeated recordings, suggested that the system should be recalibrated on a regular basis. As can be seen in Table 3, the difference between positions of sensors moving in the circal grew progressively on each recording day. This gradual increase in error might be

related to differences in environmental factors (e.g., room temperature) between different days. Thus, we recommend that the system be calibrated before each experiment in order to ensure the quality of acquired data. In the majority of cases, selected sensors will require more than one calibration (e.g., for our system, Sensors 9–12 tend to be more difficult to calibrate than the other eight sensors).

The second step in minimizing the likelihood of large recording errors is to optimize the location of a participant within the recording region of the EMA cube. We recommend that the sensors attached to the tongue and lips be positioned as close to the cube center as possible because error in this region is significantly smaller than in the peripheral regions of the recording field. Ensuring good tracking of tongue and lip movements is important because identifying errors in these data is exceptionally difficult as compared with the data from sensors attached to rigid objects such as the jaw and head.

The third step in ensuring the quality of the acquired data is to systematically check data for errors. For the sensors attached to the rigid structures, distance functions between sensor pairs can be easily calculated for the entire record. Regions where the distance between these sensors are notably different from the expected values should be eliminated from data analyses. This type of evaluation is impossible for sensors attached to the tongue and lips. Other techniques must be developed to distinguish errors in positional tracking from normal variability in speech-related movements. For example, Hoole and Zierdt (2006) briefly outlined a procedure for amplitude correction using a predictable component of the residual RMS and predicted velocities. More work of this kind needs to be made available for AG500 users.

The findings from this investigation suggest that the AG500 can be used for registering movements of the articulators during speech. However, specific steps must be taken before, during, and after acquisition to ensure accuracy of the obtained data. Additionally, a number of issues need to be addressed in the future development of the device. For example, steps need to be taken to accommodate for nonuniformity of error within the measurement field. Kaburagi and colleagues (2005) commented on potential limitations of the dipole model representing the magnetic field of the cube and suggested a multivariate B-spline model, which seemed to account for location-dependent fluctuations in the field better than the dipole model. A different calibration device/routine that allows calibration in the larger region than is currently covered by the circal might allow more accurate estimation of parameters of the magnetic field function (see Zierdt, 2007, for additional suggestions). Furthermore, a device such as the accuracy checker can be helpful in identifying predictable regions of the cube where the error tends to be consistently high. These regions should be avoided during experimental recordings by carefully positioning a participant within the recording field. Input from different laboratories about users' experiences with the system is essential to the manufacturer's updates and upgrades of the device. Progress in the development of the procedures for data acquisition and postprocessing can be expedited if various laboratories begin to share their knowledge and techniques using various available venues, one being the AGwiki Web site.

## Acknowledgments

## References

Chen HY, Murdoch BE, Goozee JV, Scott D. Physiologic development of tongue–jaw coordination from childhood to adulthood. Journal of Speech, Language, and Hearing Research 2007;50:352–360.

Green JR, Moore CA, Higashikawa M, Steeve RW. The physiologic development of speech motor control: Lip and jaw coordination. Journal of Speech, Language, and Hearing Research 2000;43:239–255.

Green JR, Moore CA, Ruark JL, Rodda PR, Morvee W, VanWitzenburg M. Development of chewing in children from 12 to 48 months: Longitudinal study of EMG patterns. Journal of Neurophysiology 1997;77:2704–2716. [PubMed: 9163386]

Hertrich I, Ackermann H. Lip-jaw and tongue-jaw coordination during rate-controlled syllable repetitions. The Journal of the Acoustical Society of America 2000;107:2236–2247. [PubMed: 10790049]

Hoole P. Issues in the acquisition, processing, reduction, and parameterization of articulographic data. Instituts für Phonetik und Sprachliche Kommunikation, München (FIPKM) 1996;34:158–173.

Hoole, P.; Zierdt, A.; Geng, C. Beyond 2D in articulatory data acquisition and analysis. In: Solé, M-J.; Recasens, D.; Romero, J., editors. Proceedings of the XVth International Congress of Phonetic Sciences; Barcelona, Spain. Aug. 2003 p. 265-268.

Hoole P, Zierdt A. Five-dimensional articulography. Stem-, Spraak-en Taalpathologie June;2006 14 (Suppl.):57.

Honda M, Kaburagi T. Comparison of electromagnetic and ultrasonic techniques for monitoring tongue motion. Instituts für Phonetik und Sprachliche Kommunikation, München (FIPKM) 1993;31:121–136.

Kaburagi T, Wakamiya K, Honda M. Three-dimensional electromagnetic articulography: A measurement principle. The Journal of the Acoustical Society of America 2005;118:428–443. [PubMed: 16119363]

Perkell J, Cohen M, Svirsky M, Matthies M, Garabieta I, Jackson M. Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. The Journal of Acoustical Society of America 1992;92:3078–3096.

Steele CM, Van Lieshout PHHM. The use of electromagnetic midsagittal articulography in the study of swallowing. Journal of Speech, Language, and Hearing Research 2004;47:342–352.

Tuller B, Shao S, Kelso JAS. An evaluation of an alternating magnetic field device for monitoring tongue movements. The Journal of the Acoustical Society of America 1990;88:674–679. [PubMed: 2212291]

Zierdt A. Problems of electromagnetic position transduction for a three-dimensional articulographic measurement system. Instituts für Phonetik und Sprachliche Kommunikation, München (FIPKM) 1993;31:137–141.

Zierdt, A. EMA and the crux of calibration. Proceedings of the XVIth International Congress of Phonetic Sciences; Aug. 2007 p. 593-596.

Zierdt, A.; Hoole, P.; Tillmann, HG. Development of a system for three-dimensional fleshpoint measurement of speech movements. Proceedings of the XIVth International Congress of Phonetic Sciences; Aug. 1999 p. 3

Zierdt, A.; Hoole, P.; Honda, M.; Kaburagi, T.; Tillmann, H-G. Extracting tongues from moving heads. Proceedings of the Vth Speech Production Seminar; May. 2000 p. 313-316.
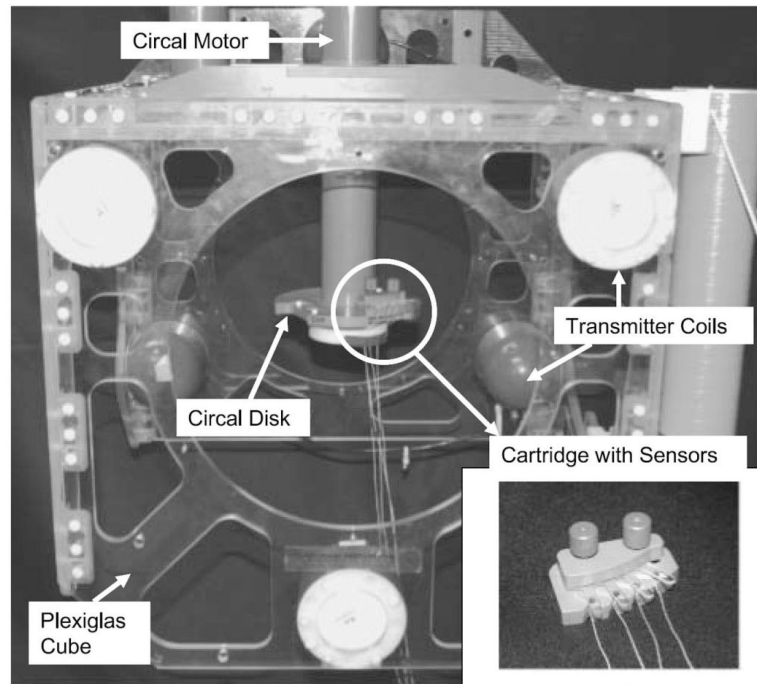
**Figure 1.**
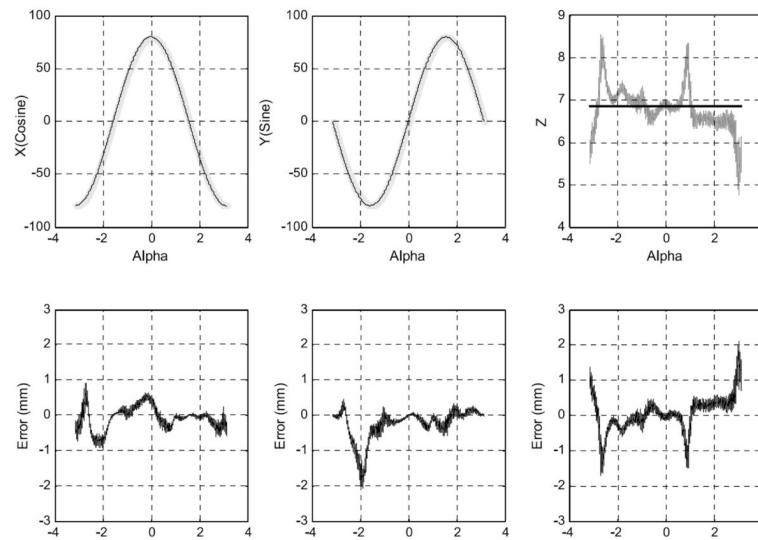The 3D electromagnetic articulography (EMA) device is shown with its parts labeled.

**Figure 2.**
Measured and expected movement histories in X, Y, and Z dimensions (top plots) and the error functions defined as the difference between the measured and expected positions (bottom plots).
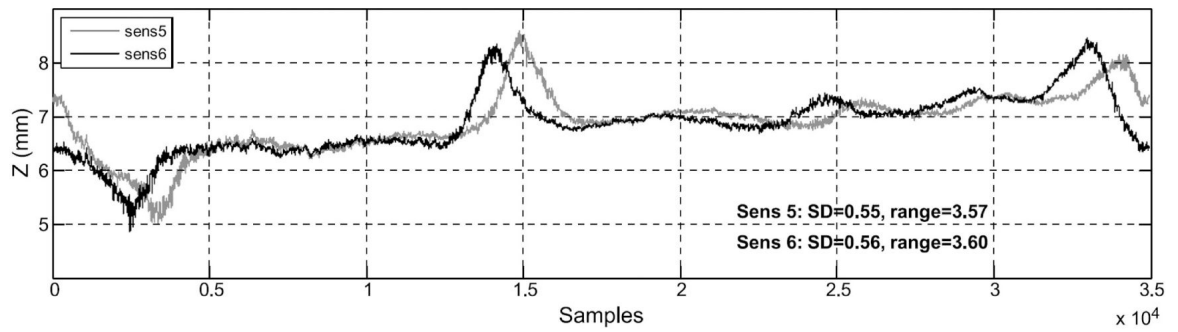
**Figure 3.**
Movement histories in the Z dimension for Sensors 5 and 6. The Z dimension was chosen for this illustration because it typically had the most observable unidimensional error in circal recordings. Note the time lag between position peaks where the largest deviations from the semistraight line occur as the sensors move around the circle in the cube. The signals seem to deviate greatly from the expected path when sensors enter specific zones in the magnetic field of the cube.
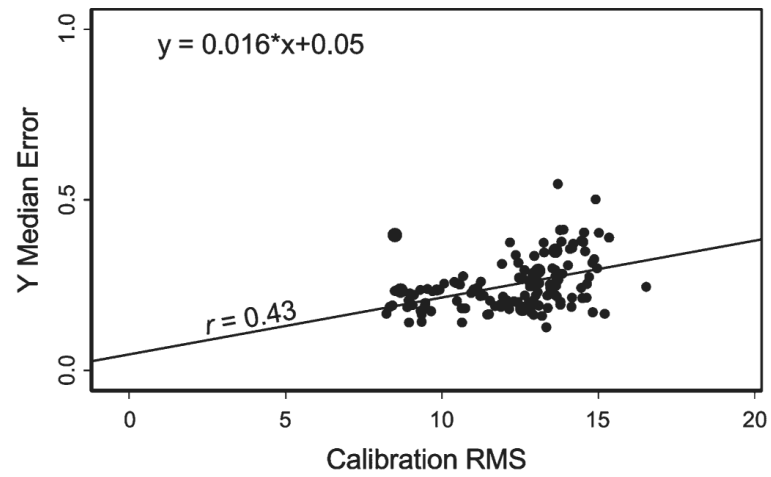
**Figure 4.**
A scatterplot showing the relationship between calibration root-mean-square (RMS) values (digits) on the *x*-axis and the Y median error (mm) on the *y*-axis.
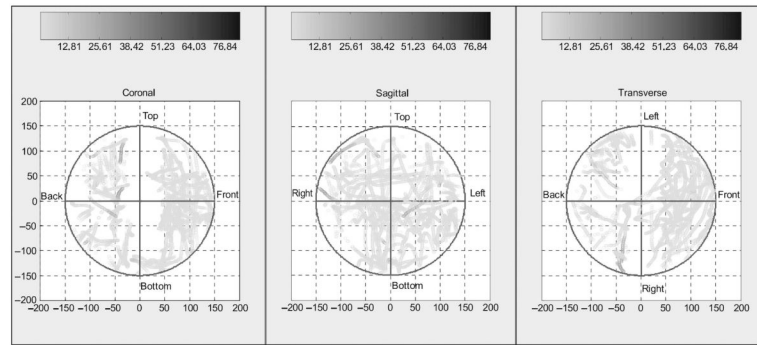
**Figure 5.**
Distribution of the spatial error defined as the mean-corrected Euclidean distance between two adjacent sensors. All errors above 0.5 mm are plotted. The optimal measurement field is defined by the circle with the radius of 15 cm. Different presentation orientations are identified on each plot.

**Table 1**

Summary statistics of the medians of absolute measurement error in X, Y, Z, R (in mm), and θ (in degrees) calculated across sensors and sessions.

| Signal | Median | IQR | Maximum |
|--------|--------|-----|---------|
| X | 0.24 | 0.42 | 1.51 |
| Y | 0.22 | 0.39 | 1.78 |
| Z | 0.38 | 0.43 | 2.53 |
| R | 0.39 | 0.52 | 1.77 |
| θ | 0.21 | 0.26 | 0.84 |

IQR = interquartile range.

**Table 2**

Summary statistics for positional errors (in mm) calculated as a distance function between two sensors attached to the jaw during four different tasks.

| Variable | Median | IQR | 95% | Maximum |
|---|---|---|---|---|
| /a/ | 0.09 | 0.06 | 0.25 | 0.35 |
| Syllables | 0.15 | 0.03 | 0.44 | 0.91 |
| Sentence | 0.16 | 0.10 | 0.61 | 1.03 |
| Paragraph | 0.22 | 0.30 | 1.72 | 2.00 |

**Table 3**

Summary statistics for the repeatability analysis computed across sensors and showing error (in mm) calculated as the difference between sensor positions recorded during circal movements on two different days.

| Comparison | Median | IQR | Maximum |
| --- | --- | --- | --- |
| Day 1/Day 2 | 0.01 | 0.15 | 0.62 |
| Day 1/Day 3 | 0.23 | 0.33 | 1.32 |
| Day 1/Day 4 | 0.30 | 0.41 | 1.93 |
| Day 1/Day 7 | 0.35 | 0.45 | 2.30 |
| Day 1/Day 9 | 0.44 | 0.58 | 2.33 |