Electromagnetic Articulography with AG500 and AG501

Massimo Stella 1, Antonio Stella 2,3, Francesco Sigona 2, Paolo Bernardini 1, Mirko Grimaldi 1, Barbara Gili Fivel 2

1 Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy
2 Centro di Ricerca Interdisciplinare sul Linguaggio (CRIL), Università del Salento, Lecce, Italy
3 Dipartimento di Studi Linguistici e Letterari, Università di Padova, Padova, Italy

massimo.stella@inbox.com, antonio.stella@unipd.it, francesco.sigona@unisalento.it, paolo.bernardini@le.infn.it, mirko.grimaldi@unisalento.it, barbara.gili@unisalento.it

Abstract

The AG500 (Carstens Medizinelektronik GmbH) is widely used in electromagnetic articulography, as it allows the recording in 3D of the movements of the articulatory organs. However some anomalies in sensor position tracking were recognized and analyzed in previous works, which ultimately attributed them to numerical issues and not to external interferences. Recently a new upgraded model, the AG501, has been introduced. By means of both speech experiments and numerical measurements, we analyzed the reliability of the AG501 and compared it to that of the previous model, the AG500. While, in some scenarios, the positions obtained via the AG500 are afflicted by external perturbations, the AG501 showed a greater degree of precision in all the numerical experiments, and in the speech analyses that have been performed so far.

Index Terms: AG501, AG500, electromagnetic articulography, speech analysis, machinery noise

1. Introduction

In the few last decades, technological development in both imaging and position tracking techniques has allowed a refined study of articulatory correlates of speech production. Electromagnetic articulography (EMA) falls within the second category of methodologies, and exploits the physical properties of electromagnetic induction to track the position of the articulators over time.

Up to 2012, one of the most widely used pieces of machinery for EMA in three dimensions is the AG500, produced by Carstens Medizinelektronik (Lengern, Germany). The AG500 allows the determination of the positions and orientations of up to 12 sensors (receiver coils) glued onto articulators and reference points, moving inside a spherical registration volume with a radius of 150 mm and immersed in a superposition of six alternating magnetic fields, each generated by one transmitter coil (further design details are available in [1]).

Several recent independent studies have pointed out the presence of some pathological anomalies in the positions retrieved by the AG500. In [2], errors up to 2 mm were observed for a pair of sensors glued on the jaw dentition during some speech tasks. Additional experiments with the receiver coils rotating, with constant velocities, along a circumference in a horizontal plane centered in the registration volume showed an adequate mean error of 0.5 mm, but with extreme values reaching up to 5 mm and distributed around a median of 2.5 mm. The accuracy of the AG500 was studied also in [3], where deviations above 1.8 mm were found in unconstrained movement of sensor coils inside the registration volume. Thus, problems concerning position retrieval are well-known and some a posteriori numerical methods have also been proposed in [4] to minimize the influence of the perturbations afflicting the computed positions and the velocities of the sensors. However, only very recently the ultimate cause of these critical instabilities has been analyzed in [5], testing the hypotheses of [2], [3] and [4] by means of constrained movement measurements, stability trials and numerical tests. Results showed that errors in position tracking are due to numerical issues and cannot be attributed to external interference.

Lately, a newer model of EMA, the AG501, has been released. The system's transmitter coils spatial distribution has been completely redesigned and shows many other improvements. It records the positions and the orientations of up to 24 HQ220-L120-B sensors with a 250 Hz operating sampling rate. Compared to the previous model, the number of transmitter coils has been increased from 6 to 9; furthermore, the coils stand above the registration volume, and the plastic cube structure which delimited it in AG500 has been eliminated.

To the best of our current knowledge, there is no publication dealing with the AG501 yet, so that this work is the first analysis performed by an independent research group of the performance of the new model of the Carstens' Articulograph. In the present paper, we compare the performance of the AG500 to that of the AG501, by basically replicating for AG501 the study described in [5] (with minor, necessary adjustments), and by focusing on speech production scenarios. We also report on the possible stable configurations for AG500 and analyze the overall accuracy of AG501. Thus, methods and results found in [5] will be summarized in the next section, as they are the starting point of the present work.

2. Checking stability of AG500

2.1. Method for accuracy assessment

In [5] we studied the accuracy of the AG500, empirically quantifying its positional errors in the tracking procedure of sensors movements, according to specific setups. Mutual inductions and external electromagnetic noise were identified as not significant in affecting simple circular movements of the receiver coils, along horizontal circumferences at different heights, in differently rotated configurations. These very same trials highlighted also some directions in which almost every reconstructed sensor position (of the nine available) was critically noisy, with standard deviations on its height ranging from 0.2 mm up to 3.9 mm. In order to completely remove any possible physical influence on the sensor localization (i.e. non-directional background noise), numerical simulations were conducted, where the AG500
software routine was fed with a fake input file with 100,000 identical entries. Each entry corresponded to the measured amplitudes for two sensors, one with coordinates afflicted by critical noise (noisy configuration) and the other with accurate coordinates (working condition). On the one hand, the output of the numerical algorithm for the heights of the sensor in the working condition showed a Gaussian curve with a standard deviation of just 0.2 mm, proving that the AG500 used during the experiments was working properly. On the other hand, the results for the receiver coil in the noisy configuration showed a bimodal distribution, with two peaks of height at 30.5 mm and 37.5 mm. This result is clear evidence that the instabilities in the AG500, recognized also by [2], [3] and [4], are to be ultimately ascribed to numerical issues, regarding convergence to the desired computed position of the Newton-Raphson algorithm, implemented in the CalcPos software offered together with AG500 by the manufacturer.

From a speech analysis point of view, the aforementioned disturbances make the study of articulatory movements rather difficult, given the fact that the experimenter cannot always recognize and distinguish noisy instabilities from minimal movements of the articulators actually occurring during speech production.

2.2. Speech measurements with AG500

The circular movement measurement performed in [5] showed that the accuracy of the AG500 in the reconstruction of trajectories is not uniform in measurement volume. Indeed, in some parts of the volume, accuracy decreases, leading to a high instability of the sensor trajectory. In order to evaluate how this impacts on the work done by the experimenters in collecting speech materials, we acquired speech data in different parts of the measurement volume. This was done by recording speech produced by a subject whose position was rotated anticlockwise in four steps of 45°. In each position - to start with the first one, in which she was parallel to the positive values of the $x$ axis - the subject repeated the speech corpus. Figure 1 shows the orientation of the head of the subject in the different positions.

The data show a different number of errors in the trajectory of the articulators, depending on position in the cube. The highest accuracy of trajectories is found in position 2, where those of all the relevant articulators involved in the speech productions are accurately detected (see Figure 2, top panel). On the contrary, the lowest accuracy is found in position 4, where the trajectories show an interference of noise, in particular for gestures which do not involve wide displacement values, such as those produced for unaccented syllables. Furthermore, the amount of errors seems to propagate progressively to other sensors leading to instable trajectories for all the sensors, as shown in the last production of the triplet (Figure 2, bottom panel).

Figure 1: Orientation of the head of the subject in the five positions with respect to the EMA cube.

The corpus used in this experiment consists of the 3 pseudo-words [mi.'ma.mi.ma], [ni.'na.na.na] and [li.'la.li.la], widely exploited in research projects conducted at CRIL (see [6], inter alia). For each position, the subject was asked to produce 3 consequent repetitions of each word. In order to maximize the displacement of the articulatory gestures involved in the production of the stressed syllable, a strong focalization was induced on the target words, by asking the subject to produce them as if they were corrections to previous statements, i.e. producing a contrastive-corrective pitch accent. The sensors were glued onto the articulators of the subject in the following order: 3 on the tongue (root, blade and tip), 2 on upper and lower lips, and 2 on upper and lower incisors.

The data show a different number of errors in the trajectory of the articulators, depending on position in the cube. The highest accuracy of trajectories is found in position 2, where those of all the relevant articulators involved in the speech productions are accurately detected (see Figure 2, top panel). On the contrary, the lowest accuracy is found in position 4, where the trajectories show an interference of noise, in particular for gestures which do not involve wide displacement values, such as those produced for unaccented syllables. Furthermore, the amount of errors seems to propagate progressively to other sensors leading to instable trajectories for all the sensors, as shown in the last production of the triplet (Figure 2, bottom panel).

Figure 2: Trajectories of 4 sensors for 3 consecutive [ni.'na.na.na] productions. Shaded areas highlight errors.

Figure 3: Displacement and velocity tracks of the tongue tip sensor for the third [ni.'na.na.na] production. Top panel: position 2; bottom panel: position 4.
Looking in detail at the trajectories of the tongue tip sensor for the last production of the triplet in positions 2 and 4 (see Figure 3), it is clear that the only detectable gestures in position 4 are the opening and closing gestures produced for [a] vowels, even if a perturbation is shown in the velocity trajectories of the gestures, leading to doubts about the reliability of the target position. Indeed, a noise seems to be superimposed to the actual velocity profile of the whole token. This does not allow the experimenter to isolate the single gestures of [nin] sequences, probably because the velocity values are lower than the noise produced by the machine. In such a case, the token should be excluded and the recording should be repeated.

3. Checking stability of AG501

3.1. Method for accuracy assessment

Exactly like the AG500, the AG501 is supplied with a plastic rotating support called circal. With a setup similar to the one used in [5], we fixed 16 sensors to a static support firmly attached to the circal, using the same orientation for all the sensors. Then, we performed circular movements of the circal at fixed speeds along circumferences of fixed height. The results of a reconstruction procedure are reported in Figure 4.

Figure 4: Top view of the reconstructed 3D positions of the rotating sensors inside the spherical registration area V (gray overlay) with diameter 300 mm.

Contrary to that which we did with the AG500, we repeated similar trials for horizontal circumferences scattered in the upper, the middle and the lower part of the spherical registration volume V, in order to cover and scan it systematically.

In this Section, we will indicate the coordinates of a sensor with \( x, y, z \), while the term radius (\( r \)) will refer to the Euclidean distance of a receiver coil from the center of \( V \), including also the contribute of the height \( z \).

In the equatorial zone, no evident peaks of Root Mean Square (RMS) were observed, the latter being a parameter, named by the manufacturer and measured in digits, that quantifies the discrepancies between the measured and the estimated binary induced amplitudes. The computed standard deviations (measured in millimeters and indicated with \( std \)) for the sensors moving inside \( V \) showed peaks over 1.0 mm, especially near the internal border of the measurement volume. The standard deviations for the receiver coils with \( r \leq 110 \) mm were upper bounded by a threshold of 0.3 mm, as shown in Figure 5. This shows the good performances of AG501, which operates according to the dynamical accuracy prescription of 0.3 mm \( std \) over displacements of 10 cm stated by the manufacturer for linear movements.

Figure 5: Differences between \( std \) (▲) and \( std \) (■), for sensors displaced in the equatorial zone of \( V \), at different distances from its center (radiuses).

Inside the middle zone of \( V \), the machine gave rise to tracked positions which were not only precise (i.e. with low values of \( std \)), but also very accurate, as confirmed by elementary fittings done with the expected circumferential patterns but not reported here for reasons of brevity.

In the lower part of \( V \) this high degree of accuracy lowered to some extent. As shown in Figure 6, the computed trajectories were subject to some noise, especially at lower heights, even if the AG501 remained fairly precise, with \( std \) values still below the 0.3 mm threshold (data not shown for space issues).

Figure 6: Reconstructed circular trajectories with heights \( z_l \approx -122.5 \) mm (left) and at height \( z_m \approx -52.3 \) mm (right) and ideal circumferences, dashed.

An increase in both precision and accuracy was measured in the upper part of \( V \), where all the standard deviations on the sensors are below the threshold value of 0.3 mm, once again corroborating the manufacturer's guarantee of accuracy regarding efficient reconstruction of sensor movements.

Figure 7: Differences between \( std \) (▲) and \( std \) (■), for sensors displaced in the upper zone of \( V \), at different distances from its center (radiuses).

As reported in Figures 5 and 7, increasing the distance of the sensors from the center of the registration volume leads the
dynamical accuracy of the $z$ coordinate to decrease much less than that of the radius (which contains also the contribution of $x$ and $y$). These findings suggest that the height is a rather stable coordinate, whose dynamical accuracy is not dependent on the receiver coil positions in the upper and middle parts of $V$. Therefore, it could be used as an additional parameter to be checked in speech production experiments performed via the AG501. However, further analyses of this matter and of any possible influences of orientation are in progress.

3.2. Speech measurements with AG501

Recordings were performed similarly to those described in section 2.2 for the AG500, in exactly the same environment, with the same experimental subject and the same experimental task. In this section we offer both a discussion of results of speech recordings and some qualitative considerations on the use of the AG501 in comparison to the AG500.

As for our qualitative observations, first of all, the AG501 appears to be more stable than the AG500 from an electrical point of view. As far as the set-up is concerned, apart from the calibration phase and the phase corresponding to the preparation and gluing of the sensors, which are similar to that needed for the AG500, the other phases are more user-friendly and stable in the AG501 than in the AG500. For instance, the power-up is shorter (15 min vs. at least 2 hours), the bootstrap execution is much simpler, as well as the data export at the end of the recording session.

According to our preliminary analyses of recorded speech, data acquired by means of the AG501 showed a lower number of errors in the trajectory of the articulators, with RMS generally under the accuracy threshold of 20 digits, as prescribed by the manufacturer. Furthermore, there are no evident differences arising from the subject position with respect to the transmitter coils. Recorded data globally show a much lower amount of perturbation. In AG501 recordings, no abnormal tracking of sensors was detected in speech trajectories on the $z$ axis in the 5 positions monitored. In Figure 8, the trajectories acquired for the entire corpus in position 2 and 4 are reported. As it is clear from the figure, in both positions no perturbations are found and all the expected differences in gestural composition of the words [mi.'ma.mi.ma], [ni.'na.na.na] and [li.'la.li.la] are correctly detected (the same holds true also for the positions 1, 3 and 5). Position 4, which was the most unstable position in AG500 recordings, shows high reliable trajectories in AG501 ones.

Comparing the tongue tip trajectories of [ni.'na.na.na] recorded with AG500 in position 2, i.e. the most stable position (Figure 3, top panel), to those recorded with AG501 in the same position (Figure 9), no difference in reliability are found in either displacement or velocity dimensions. Indeed, the productions show that the same gestural pattern and the differences in displacement and velocity between the two tokens should be ascribed to the dynamical variability of the articulatory dimension of speech. The great similarity of the plots reported for the two tokens shows that, from a gestural point of view, data offered by the AG500 in position 2 could be considered reliable as those acquired by means of AG501. Nonetheless, further investigations on the $x$ and $y$ axes are required, in order to compare speech data that may be relevant in the investigation of a wider range of linguistic events.

Figure 8: Trajectories of four sensors for 3 consecutive productions of [mi.'ma.mi.ma], [ni.'na.na.na], and [li.'la.li.la]. Top panel: position 2; bottom panel: position 4.

Figure 9: Displacement and velocity tracks of the tongue tip sensor for the third [ni.'na.na.na] production in position 2.

4. Conclusions

Our findings provide an ulterior assessment on the overall high degree of accuracy of the AG500, which however fails in some situations, mainly because of numerical issues [5]. Some configurations which can greatly reduce this risk do exist, e.g. the definition of stable areas of the measurement volume, which are useful to know for speech production experiments. In general, however, the accuracy of the AG501 is by far superior, especially when measurements are in the upper zone of the volume, although further trials are necessary in order to completely rule out numerical issues. Also in the speech production analysis, the AG501 provides a greater deal of accuracy and is more user-friendly than its predecessor.

5. Acknowledgements

We would like to thank Sonia D’Apolito, who kindly offered to support this research as subject in the experimental sessions with AG500 and AG501.
6. References


