

A MATLAB toolbox for the analysis of articulatory data in the production of speech

NOËL NGUYEN

Université de Provence, Aix-en-Provence, France

The goal of this paper is to present EMATOOLS, a set of scripts for displaying and annotating acoustic and articulatory data simultaneously in studies on speech production. These scripts were developed with the use of MATLAB, a multiplatform computing environment for numeric computation and visualization. The system is equipped with a mouse-driven graphical interface made up of a number of displays. This interface can be easily customized to speed up routine tasks. The scripts can also be used in a noninteractive way, as stand-alone MATLAB commands. Output data can be imported into any standard spreadsheet. EMATOOLS is freely available from www.lpl.univ-aix.fr/~nguyen/ematools.html.

EMATOOLS is a signal editor for displaying, annotating, and processing articulatory and acoustic data recorded in studies on speech production. The system has been more particularly designed to be used with EMA and EPG data, although it is suitable for analyzing any kind of kinematic data recorded using point source tracking devices (such as the X-Ray Microbeam facility, or the OPTOTRACK optical system, among others).

EMA (electromagnetic articulography) is a technique that allows articulatory movements to be monitored by means of small electromagnetic coils attached to vocal tract structures in the midsagittal plane (Hoole, 1993; Hoole & Nguyen, 1999; Perkell et al., 1992). Measurement points can be the tongue, the upper and lower lips, the mandible, and the velum. In addition, coils are generally attached to fixed structures such as the bridge of the nose and the upper central incisors to provide a maxillary frame of reference. Alternating magnetic fields generated by transmitter coils make it possible to measure the position of each receiver coil relative to two orthogonal axes (referred to as the *x*- and *y*-axes throughout), in the midsagittal plane, with a measurement bandwidth typically ranging from dc up to about 250 Hz. The reader is referred to Perkell et al. for a detailed description of EMA. Although commercially available EMA systems are relatively recent, the number of research centers equipped with such a facility is increasing rapidly (see www.cstr.ed.ac.uk/artic for a regularly updated list of EMA users).

EPG (electropalatography) is a technique for recording the timing and location of tongue-palate contact patterns during speech (Hardcastle, Jones, Knight, Trudgeman, & Calder, 1989). It involves the subject's wearing an artificial plate molded to fit the upper palate and containing a number of electrodes mounted on the surface to detect lingual contacts. The Reading EPG2 system for example, which was used for the purposes of the present paper, employs 62 electrodes arranged in eight horizontal rows across the palate. The Kay Elemetrics Palatometer (perhaps more common in the US) uses a palate embedded with 96 electrodes. EPG patterns are usually digitized at a sampling frequency ranging from 100 to 200 Hz. EPG is now well established in many experimental laboratories and speech therapy clinics throughout the world.

When combined, EMA and EPG provide us with complementary pieces of information about the tongue movements in two orthogonal planes (the midsagittal and coronal planes, respectively). These techniques both have the advantage of being noninvasive, enabling the experimenter to collect large amounts of data for a significant number of speakers.

Our signal editor was developed with the use of the MATLAB programming language, a commercial product from The MathWorks, Inc. MATLAB is a technical computing environment for numeric computation and visualization. It integrates numerical analysis, matrix computation, signal processing, and graphics in the same environment. One of the advantages of MATLAB is that versions of the system exist for most major modern computer platforms (including Unix, Windows, and Macintosh; see www.mathworks.com for an exhaustive list of currently supported platforms). MATLAB scripts are simply made up of lists of instructions stored in ASCII files and can be run without modifications on different machines.

Admittedly, there are already quite a few excellent signal processing packages available for analyzing acoustic and articulatory data.¹ Particularly relevant to the pres-

The data shown in Figure 1 were recorded at the Institute for Phonetics of the University of Munich by Phil Hoole, Fiona Gibbon, and Bill Hardcastle. Hélène Loevenbruck and Dorothée Lefèbvre provided the scripts for importing Kay EPG/audio data into EMATOOLS. Thanks are also due Janet Fletcher, Fiona Gibbon, Justine Goozée, Bill Hardcastle, Katerina Nicolaidis, Daniel Recasens, Mark Tiede, Alice Turk, Béatrice Vaxelaire, Paula West, and Alan Wrench, for useful suggestions. Correspondence should be addressed to Noël Nguyen, Laboratoire Parole et Langage, Université de Provence, 29 avenue Robert Schuman, 13621 Aix-en-Provence, France (e-mail: noel.nguyen@lpl.univ-aix.fr).

ent paper are HADES, EMU, and MAVIS.² EMATOOLS was specifically designed with a view to displaying and analyzing EMA and EPG data, while taking full advantage of MATLAB's numerical and graphical capabilities.

OVERVIEW OF THE SYSTEM

The Graphical Interface

EMATOOLS graphical interface is made up of a number of windows each associated with a particular type of display, like those shown in Figure 1. The Main window acts as a parent environment for all functions and allows a set of acoustic and articulatory signals to be edited and

annotated. The maximal number of signals simultaneously displayed in this window depends mainly on the size of the screen.

Most of the usual editing functions (zooming, annotating, etc.) are performed with the use of two cursors, which can be displaced along the time axis with the mouse. Markers can be placed anywhere in a signal and be stored in ASCII files for later use. It is also possible to measure the sample value at the location of a cursor for a given signal and save this value to a file, which can then be imported into any standard spreadsheet for further processing. In order to make the annotation of signals a little less tedious, the system also enables specific articulatory

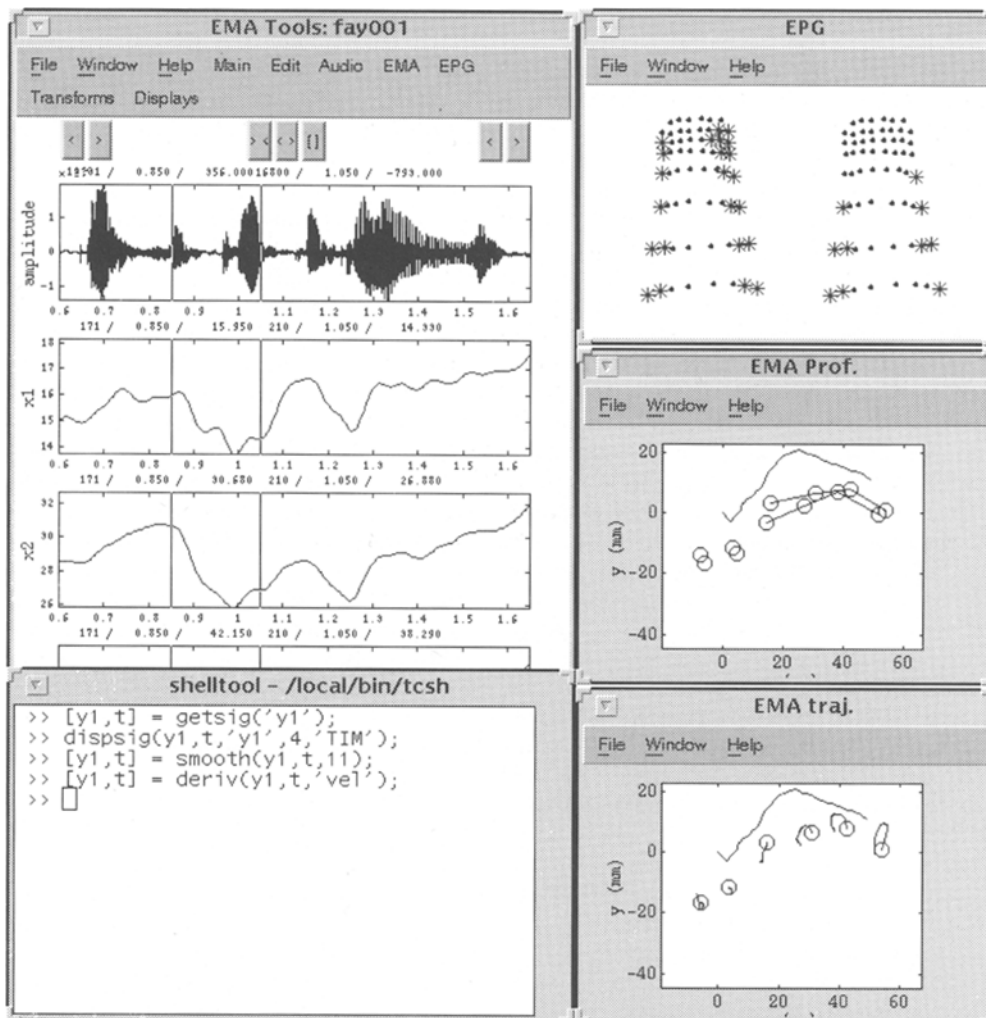


Figure 1. Some of the displays available in EMATOOLS. Top left: the Main window. Top right: EPG patterns at the location of each of the two cursors along the time axis in the Main window. Center right: positions of the coils in the midsagittal plane at the location of each of the two cursors. Bottom right: trajectories of the coils in the midsagittal plane over the time interval between the two cursors. Bottom left: the MATLAB command window.

events to be located in an automatic manner (e.g., local extrema on EMA curves, EPG patterns showing a closure in a particular region of the palate, etc.).

The signals that can be displayed in the Main window are of four different kinds: (1) the acoustic signal; (2) the EMA parameters available (x - and y -coordinates of each coil); (3) the EPG indices most commonly used in the literature (center of gravity, sums of contacts in various regions on the electropalate, lateral asymmetry, etc.); and (4) a set of parameters derived from those mentioned above by means of various transformations—for example, velocity or acceleration of the displacements of a given coil relative to one of the two axes of the coordinate system, or to both of them.

From the Main window, it is possible to access a number of other displays, some of which are shown in Figure 1. The EPG display (top right) shows the spatial arrangement of the tongue–palate contacts at the time location of each of the two cursors in the utterance (in this example, the sentence *It's a kip again*, recorded by a female speaker of English). Anterior is on top. The EMA Profile display (center right) shows the positions of five receiver coils in the midsagittal plane, at the same two time locations. Anterior is to the left. Coils were placed (from front to back) on the lower lip, the lower central incisors, the tip of the tongue, the blade of the tongue, and the anterior and posterior parts of the tongue dorsum. Points associated with tongue coils are connected with each other in the display. An outline of the hard palate is also shown. Both the EMA Profile and the EPG displays are dynamically linked to the Main window; they are permanently refreshed as cursors are moved along the time axis in this window. The EMA Trajectory display (bottom right) shows the trajectories of the EMA coils in the midsagittal plane over the time interval between the two cursors. The origin of the trajectory for each coil is indicated by a circle.

Customization, Macros, and Stand-Alone Functions

In designing EMATOOLS, we attempted to associate each function with a particular MATLAB script, or M-file, whenever this was possible. As a result, most of these functions can be performed by typing the name of the corresponding command in the MATLAB command window, therefore bypassing the pop-up menus of the graphical interface. These commands can themselves be stored in a new M-file, which can then be made available to the user in exactly the same way as those already existing—that is, either noninteractively, from the command window, or by being integrated into the graphical interface as a new, user-specific menu. This is a convenient method for creating macros to streamline repetitive tasks.

By way of example, Appendix A contains a set of commands allowing the user (1) to display a curve referred to as y_1 , which represents variations in the position of the tip

of the tongue along the y -axis, (2) to smooth this curve using a cubic spline function, and (3) to compute the first derivative (velocity) of the smoothed curve.

Processing Large Amounts of Data

EMA and EPG both allow large amounts of data to be recorded and, as a result, make it necessary to devise efficient data processing and data reduction tools. For example, extracting articulatory data at a particular point in an utterance can be quite tedious for repeated measurements of tokens or speakers. Although EMA- and/or EPG-related measures can be extracted interactively with EMATOOLS, it is also possible to do this faster and on a larger scale by means of a set of scripts that were developed as an extension to the toolbox. These scripts make use of the many advantages of the MATLAB language for processing large data sets. For example, Appendix B shows how to compute and display the average position of each of the EMA coils in the midsagittal plane from a set of x - and y -coordinates as contained in an input EMA data file. Further details and illustrations are provided in the EMATOOLS user's manual.

CONCLUSIONS AND FUTURE DEVELOPMENTS

In this paper, we have presented a new set of programs for editing, annotating, and processing acoustic, EMA, and EPG data in an integrated environment. Although it is still at the experimental stage, this system has already proved to be a useful tool in studies on speech production (see, e.g., Fletcher, 1997; Goozée, Murdoch, & Theodoros, 1999; Scobbie, Turk, & Hewlett, 1998; West, 1999; Wrench, 1999). It has been developed with a view toward taking full advantage of MATLAB numerical and graphical capabilities. Further improvements will include the integration of data reduction techniques such as principal-component analyses and multilayer neural networks. EMATOOLS is available free at www.lpl.univ-aix.fr/~nguyen/ematools.html.

REFERENCES

- FLETCHER, J. (1997). Vowel to vowel coarticulation: An EMA/EPG study. In *2nd European Symposium on Electrolaryngography and Electropalatography* (pp. 23-25). Queen Margaret College, Edinburgh, Department of Speech and Language Sciences.
- GOOZÉE, J., MURDOCH, B. E., & THEODOROS, D. G. (1999). Electropalatographic assessment of articulatory timing characteristics in dysarthria following traumatic brain injury. *Journal of Medical Speech—Language Pathology*, 7, 209-222.
- HARDCASTLE, W. J., JONES, W., KNIGHT, C., TRUDGEON, A., & CALDER, G. (1989). New developments in electropalatography: A state-of-the-art report. *Clinical Linguistics & Phonetics*, 3, 1-38.
- HOOLE, P. (1993). Methodological considerations in the use of electromagnetic articulography in phonetic research. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, 31, 43-64.
- HOOLE, P., & NGUYEN, N. (1999). Electromagnetic articulography in

- coarticulation research. In W. J. Hardcastle & N. Hewlett (Eds.), *Instrumental studies of coarticulation* (pp. 260-269). Cambridge: Cambridge University Press.
- PERKELL, J. S., COHEN, M. H., SVIRSKY, M. A., MATTHIES, M. L., GARABIETA, I., & JACKSON, M. T. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, **92**, 3078-3096.
- SCOBIE, J. M., TURK, A., & HEWLETT, N. (1998, July). *The articulation of Scottish vowels*. Paper presented at the Sixth Conference on Laboratory Phonology, York, U.K.
- WEST, P. (1999). The extent of coarticulation of English liquids: An acoustic and articulatory study. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (Vol. 3, pp. 1901-1904). Berkeley: University of California Press.
- WRENCH, A. (1999). An investigation of sagittal velar movement and its correlation with lip, tongue, and jaw movement. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (Vol. 1, pp. 435-438). Berkeley: University of California Press.

NOTES

1. Information about these packages is available at svr-www.eng.cam.ac.uk/comp.speech.
2. HADES is a program for the display and analysis of acoustic and physiological signals on VAX computers, developed by Phil Rubin at Haskins Laboratories (see www.haskins.yale.edu). EMU is a program for the creation, manipulation, and analysis of speech databases for Windows, Unix, and Linux platforms, developed by Steve Cassidy and Jonathan Harrington at Macquarie University, Sydney (see www.shlrc.mq.edu/emu). MAVIS is a set of MATLAB functions for viewing articulatory data generated by point source tracking methods such as EMA, developed by Mark Tiede at ATR-HIP Laboratories (see www.hip.atr.co.jp/~tiede).

APPENDIX A

A Short Script for Computing Articulatory Velocity

```
[y1,t] = getsig('y1');           % save Parameter y1 to Matrix y1
dispsig(y1,t,'y1',4,'TIM');     % display y1 on Track #4
[y1,t] = smooth(y1,t,11);       % smooth y1
[y1,t] = deriv(y1,t,'vel');     % compute velocity
```

APPENDIX B

A Short Script for Computing the Average Positions of the EMA Coils for a Given Utterance

```
[x,t] = readsig('fay001.ema','x',200); % read x coord. of all coils for Item 1
[y,t] = readsig('fay001.ema','y',200); % read y coord. of all coils for Item 1
plot(mean(x)',mean(y)');          % display average positions of the coils in x/y plane
```

(Manuscript received June 23, 1999;
revision accepted for publication May 29, 2000.)