Measuring physiological speech entrainment with electromagnetic articulography (EMA)

Phil Hoole & Jonathan Harrington
IPS Munich
Setting up a dual EMA configuration for studying speaker interaction

Some EMA background

A set of transmitter coils generates an alternating electromagnetic field at frequencies of about 10kHz

==> Induction of a position- and orientation-dependent signal in small sensors attached to the articulators

Captures data for both visible (including head) and invisible articulators in readily analyzable form

Good temporal resolution (samplerate typically > 200Hz)
First approach: Edinburgh Speech Production Facility
Two Carstens AG500 machines
Main drawback: Both systems use the same transmitter frequencies
==> interference ==> speakers must be at least 6m apart
==> in practice, in separate rooms (connected by CCTV)

Second Approach

1x Carstens AG500 with 1x NDI Wave
No mutual interference of transmitter frequencies
Speakers can be face-to-face, about 2m apart
Pilot experiment at MARCS lab, Univ. W. Sydney:
Our setup: 1 x Carstens AG501 with 1x Carstens AG500
No mutual interference at inter-speaker distances > 2m
Detailed view of the two synch signals at trial onset
Further features of recording setup

- No noisy equipment in recording booth
- HDV video: merged view of 2 cameras plus control PC monitor
- Additional synchronized measurement signal are easy to record
Pilot experiment based on Tiede et al. experiment at MARCS

EMA sensors

articulators: tongue tip, tongue body, upper lip, lower lip
head: upper incisors, bridge of nose, behind left and right ear

After recording, head sensors used
(1) to factor out head movement from articulator sensors
(2) to calculate the rigid-body parameters of head movement (3 translations, 3 rotations)
Subjects spoke alternating word sequences for 30s, e.g.

S1: “Topf Kopf Topf Kopf ......”
S2: “Kopf Topf Kopf Topf ......”
sometimes “disturbed” by instructions from investigator to speed up or slow down.


[some non-alternating sequences also recorded, e.g.

S1: “Topf Topf Topf Topf ......”
S2: “Kopf Kopf Kopf Kopf ......”]
S1 and S2 (sagittal): tongue

S1 and S2

TT

TB

Y (mm)

Z (mm)
Movie demo
twinema1_cut: Cut 10/37. "1_topfkopf_2_kopftopf_N<AG500 10  189454>" (Type 0) 30.08 s. Trial 10

TB intrusion: S1
TT intrusion: S2

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>mm</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>mm</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>mm</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
</tr>
</tbody>
</table>
Final comments

Examples of entrainment in this presentation based on rhythmic utterances (“speech-error” paradigm).

But Tiede & Mooshammer have recently applied the dual-EMA setup to a more traditional phonetic convergence paradigm.

Kinematic properties of velar consonants in test items more similar after a period of face-to-face interaction.

Outlook

Upgrade AG500 to AG501 (2x AG501 setup is possible because AG501 can be configured to different transmitter frequencies)

Advantages of AG501:
- data-processing much less time-consuming
- large head-movements captured with much less measurement error
- subject less enclosed (easier to combine with optical systems)
  ==> improved possibilities for more natural subject interaction

Other labs (e.g. USC) are exploring use of dual NDI Wave systems.