Topographic recordings of auditory evoked potentials to speech: Subcortical and cortical responses

LUDOVIC BELLIER, PATRICK BOUCHET, ARNAUD JEANVOINE, OLIVIER VALENTIN, HUNG THAI-VAN, AND ANNE CACLIN

A Lyon Neuroscience Research Center, Brain Dynamics and Cognition team, INSERM U1028, CNRS UMR5292, Lyon, France
bUniversity Lyon 1, Lyon, France
cHospices Civils de Lyon, Service d’Audiologie et Explorations Orofaciales, Hôpital Edouard Herriot, Lyon, France

Abstract

Topographies of speech auditory brainstem response (speech ABR), a fine electrophysiological marker of speech encoding, have never been described. Yet, they could provide useful information to assess speech ABR generators and better characterize populations of interest (e.g., musicians, dyslexics). We present here a novel methodology of topographic speech ABR recording, using a 32-channel low sampling rate (5 kHz) EEG system. Quality of speech ABRs obtained with this conventional multichannel EEG system were compared to that of signals simultaneously recorded with a high sampling rate (13.3 kHz) EEG system. Correlations between speech ABRs recorded with the two systems revealed highly similar signals, without any significant difference between their signal-to-noise ratios (SNRs). Moreover, an advanced denoising method for multichannel data (denoising source separation) significantly improved SNR and allowed topography of speech ABR to be recovered.

Descriptors: Speech ABR, FFR, N1, Event-related potential, Brainstem

Speech auditory brainstem response (speech ABR) is a scalp-recorded, high frequency event-related potential (ERP) evoked by speech stimuli. As speech contains brief transitions and long periodic parts, speech ABR has two types of components: transient and sustained, respectively identified as onset/offset response and frequency-following response (FFR). With its high specificity to each speech stimulus, speech ABR is an ideal electrophysiological marker of speech encoding, and has become widely used in numerous studies assessing a large range of hearing particularities, from expertise to impairments. For instance, musicians’ ability to better hear speech in noise is related to their faster neural timing as revealed by speech ABR (Parbery-Clark, Strait, Hittner, & Kraus, 2013). And poor pitch tracking in speech ABR’s FFR component revealed by speech ABR (Parbery-Clark, Strait, Hittner, & Kraus, 2013). And poor pitch tracking in speech ABR’s FFR component.

When it comes to speech ABR recording, clinical electroencephalography (EEG) systems originally designed for click ABR measurements are most often used, as they allow for the recording of high frequency activities with fine temporal resolution and decent signal-to-noise ratios due respectively to their high sampling rate (7 to 50 kHz) and amplifier quality. However, such EEG systems have a major limit: their very low number of channels (often three including ground and reference), which constrains the experimenter to use the classical “vertical montage” (vertex referenced at the mastoid ipsilateral to the stimulated ear or at linked mastoids). As a result, the experimenter has no access to the topographical information required, for example, for speech ABR generator(s) characterization, a topic scarcely debated in the literature. Access to speech ABR topography could also allow the enhancement of the description of populations of interest.

Obtaining speech ABR topographies could be done using research EEG systems, which are classically used to record cortical auditory evoked potentials (AEPs) on a large number of channels (32 to 256), at the expense of a relatively low sampling rate (below 5 kHz). In principle, referring to the Nyquist-Shannon sampling theorem (the maximum observable frequency is half the sampling frequency) and to the brainstem phase-locking limit of 1500 Hz (reviews in Galbraith et al., 2000; Richard, Jeanvoine, Veuillet, Moulin, & Thai-Van, 2010), such a system should be able to record proper speech ABR. Moreover, almost every speech ABR study...
uses alternating polarities for stimulus presentation. This technique consists in presenting the stimulus and its opposite phase in equal amounts, then averaging speech ABRs to both polarities: this can minimize a possible stimulus artifact and the cochlear microphonics contribution to the response, yet it also emphasizes lower frequencies up to 800 Hz (Aiken & Picton, 2008). A tutorial on speech ABR recordings (Skoe & Kraus, 2010) reports that it is possible to use low sampling rate EEG systems (~6 kHz) for speech ABR recording. A few studies have indeed reported speech ABRs with sampling frequencies below 5 kHz (Aiken & Picton, 2006; Dajani, Purcell, Wong, Kunov, & Picton, 2005; Prévost, Laroche, Marcoux, & Dajani, 2013). However, none of them have compared the resulting speech ABRs with those recorded with higher sampling rate EEG systems. Controlling the impact of using a conventional research low sampling rate, multichannel EEG system on speech ABR quality is thus a mandatory step prior to further processing and analysis of such data.

To this aim, we first compared speech ABRs simultaneously recorded with two EEG systems: a 4-channel clinical EEG system with a sampling rate of 13.3 kHz, and a 32-channel research EEG system with a sampling rate of 5 kHz. Then, using data from the research EEG system only, we obtained concomitant cortical AEPs and their topographies and, for the first time to our knowledge, the topography of speech ABR, thanks to an advanced method of signal processing based on spatial denoising, allowed by the 32-channel recordings. We also show how multichannel recordings can be used to reduce acquisition time.

Method

Participants

Fourteen healthy adults (ages 20–48 years, M = 28 years; 5 women) participated in the study. All were normal-hearing (no threshold below 20 dB HL between 250 and 8000 Hz in tonal audiometry, and wave V latencies below 6.4 ms in click ABRs) native French speakers, and none had known language, psychiatric, or neurologic impairment. Each participant gave his/her informed written consent. All experimental procedures were carried out following the Declaration of Helsinki and with the agreement of the Ethics Committee from Lyon’s Leon Bérard Center (CCPRRB #05-024).

Stimulus

Speech stimulus was a 200-ms natural /ba/ syllable, recorded from a French female voice (Figure 1a). It was composed of a 108-ms voicing followed by a 92-ms vowel. Fundamental frequency of the vowel /a/ varied slightly around 200 Hz, and the four first formants were respectively centered around 825, 1490, 3070, and 4430 Hz. A 30-ms silence was inserted at the beginning of the file to enable the recording of a baseline epoch in the EEG data (see below).

Figure 1. Speech ABR evoked by a 200 ms-long /ba/ syllable (a; A.U. = arbitrary units), simultaneously recorded with clinical and research EEG systems, for two representative participants (b,c). For both participants, speech ABR from clinical and research EEG systems highly resemble each other, with similar noise level (0 to 30 ms), signal amplitude, and time course (150 to 225 ms). Note the FFR difference between S1 (b) and S12 (c), respectively, following 100 Hz (/a/ fundamental frequency subharmonic) and 200 Hz.
**Procedure**

Participants sat in an armchair in a quiet room. The stimulus was binaurally delivered through electrically shielded insert earphones (Akhoun et al., 2008), at a comfortable intensity of 50 dB SL (sensation level), with a rate of 1.1 per second, using a MATLAB (The MathWorks, USA) script to randomize stimulus polarities (1,500 presentations of each polarity). Participants were asked to remain awake during the recordings and count the number of times a green diode in front of them turned red (5% of the trials). The AEP recordings lasted 45 min with two breaks at 15 and 30 min.

**Electrophysiological Recordings**

Electrophysiological activity in response to the /ba/ syllable was recorded simultaneously by two EEG systems. A custom-made electronic device ensured that both systems were triggered simultaneously. The clinical Centor USB system (Racia-Alvar, France) recorded EEG signals from two derivations (sampling rate: 13.3 kHz; 70 Hz high-pass filter), both using the vertex (Cz) as the active electrode and one of the mastoids as the reference electrode (ground on the nasion), during 300-ms-long epochs triggered by the beginning of the sound file. The research BrainAmp system (Brain Products, Germany) recorded the EEG signal continuously (sampling rate: 5 kHz; 0.1 Hz high-pass filter) from 31 electrodes positioned in a cap following the 10–20 system, and one electrooculogram electrode at the outer canthus of the left eye—all 32 electrodes referenced to the nose (ground on the right cheek). Both systems used passive Ag/AgCl electrodes (impedance below 5 kΩ). Vertex and mastoid electrodes for both systems were positioned less than 1 cm away from each other.

**Comparison of Speech ABRs Obtained With the Two EEG Systems**

Data were analyzed using ELAN (Aguera, Jerbi, Caclin, & Bertrand, 2011) and MATLAB. For each participant, raw data from both EEG systems were first filtered with notch filters to remove the 50 Hz power line radiation artifact and its odd harmonics, then with an 80–1000 Hz Butterworth band-pass filter (fifth order) to isolate subcortical activity (additionally, to obtain cortical AEP further using information on the 28 other electrodes. The signals from the two systems were compared in two montages. The signals from the two systems were compared in two montages. The clinical Centor USB system (Racia-Alvar, France) recorded EEG signals from two derivations (sampling rate: 13.3 kHz; 70 Hz high-pass filter), both using the vertex (Cz) as the active electrode and one of the mastoids as the reference electrode (ground on the nasion), during 300-ms-long epochs triggered by the beginning of the sound file. The research BrainAmp system (Brain Products, Germany) recorded the EEG signal continuously (sampling rate: 5 kHz; 0.1 Hz high-pass filter) from 31 electrodes positioned in a cap following the 10–20 system, and one electrooculogram electrode at the outer canthus of the left eye—all 32 electrodes referenced to the nose (ground on the right cheek). Both systems used passive Ag/AgCl electrodes (impedance below 5 kΩ). Vertex and mastoid electrodes for both systems were positioned less than 1 cm away from each other.

Speech ABRs evoked by the /ba/ stimulus (Figure 1a) are shown in Figures 1b,c for two participants. Visual inspection of the curves suggests similar noise levels and time courses of signal for both systems. Expected components of speech ABR are retrieved with the two systems. A statistical analysis of the ABR components revealed that the two EEG systems had different latency distributions and signal-to-noise ratios. The ABRs were obtained with the two systems were found at a constant lag of around 0.75 ms, which may be the reflection of online filter differences between the two EEG systems. We found an average r² of .39 (Pearson correlation coefficient, see Table 1; the correlation for each participant was significant, p < .0001), showing high resemblance between the two signals as they share almost 40% of their variance. Moreover, comparison of SNRs between the two EEG systems showed no significant difference (paired t test: t(13) = 1.448, p = .17). Altogether, visual inspection of the temporal courses of ERPs, cross-correlation, and SNR measurements for the two EEG systems support the view that it is possible to record speech ABR with a classic multichannel research EEG system.

As a consequence, it is possible to benefit from all the information provided by such a system: using information on the 31 channels for advanced processing or analysis methods, but also getting cortical AEP and their topography (Figure 2d), and for the first time topographical information on speech ABR. Figure 2b,c, respectively, shows grand averages of speech ABRs (N = 14) observed at Cz rereferenced to the average of the mastoids (classical vertical montage) and speech ABR obtained with DSS (allowing us to obtain an optimal montage corresponding to speech ABR topography), both evoked by the /ba/ stimulus shown in Figure 2a. Speech ABR topography (Figure 2c) remarkably emphasizes the usefulness of the classical vertical montage (vertex–mastoids), but goes further using information on the 28 other electrodes.
The multichannel denoising method tested here, DSS, obviously increases SNR as confirmed in Table 1 (paired t test: \(t(13) = 3.788, p = .002\)). Furthermore, as shown in Figure 2e, the SNR obtained using a vertical montage by averaging all acceptable trials (mean 1,672 across participants) from the research EEG system is reached and even exceeded in only 1,000 trials using DSS. Using DSS on the 500 first trials led to speech ABRs with an acceptable SNR of almost 3 dB (Figure 2e), as is the case for the clinical EEG system.

### Discussion

The present article reports a new methodology for multichannel recording of speech ABR with a critical comparison of the same signal recorded with two different EEG systems; to the best of our knowledge, this is the first control of the effect of lowering sampling rate on speech ABR’s quality. We show here that it is possible to record high-quality speech ABR using a research EEG system with a relatively low sampling rate (here, 5 kHz). The speech ABRs thus recorded were indistinguishable in terms of SNR from speech ABRs acquired with a clinical EEG system as is most often done in speech ABR studies. Note that SNRs in this study were comparable and even superior on average to SNRs measured in other studies, which range typically between 2.5 and 3 dB (Skoe & Kraus, 2010), whatever the EEG system or data processing method used, thus confirming the validity of our results.

We retrieved here a similar mean correlation coefficient between the speech ABRs recorded with a clinical and a research EEG system as in a within-session test-retest reliability study of speech ABR (Hornickel, Knowles, & Kraus, 2012; note, however, that 8- to 13-year-old children were tested in this previous report). In the test-retest study, the correlation coefficient between two speech ABRs from two separate runs of 3,000 trials in the same session was on average 0.71 (Spearman’s rho), with a standard deviation (SD) of 0.3. Here, mean Spearman’s rho across our 14 participants is 0.59 (SD = 0.17). The small difference could be explained by the different positions (1 cm) between identical electrodes of both EEG systems, and by the slightly lower number of trials in the present study.

A variant of alternating polarities, consisting of subtracting speech ABR to both polarities instead of summing them, put the emphasis on the high frequency components of the signal, without minimizing possible stimulus artifact or cochlear microphonics contribution to the response. Using this variant (Aiken & Picton, 2008), there might be an advantage to record speech ABR with a very high sampling rate EEG system, but as the large majority of speech ABR studies uses the classical alternating polarities approach (summation), the interest of our methodology is still valid.

Thus, the present methodology enhances classical speech ABR recording, by allowing the simultaneous collection of cortical AEP and speech ABR topographies. In other words, it provides a wide panorama of electrophysiological activities evoked by a speech stimulus within a single recording session. Moreover, topographies of speech ABR have never been described, although they may represent an interesting electrophysiological marker in addition to latency and amplitude measurements, and allow the experimenter to gain insight into the poorly known generator(s) of this response. Note that a few reports of click ABR topography can be found in the literature (e.g., Norrix & Glatkic, 1996). Using this methodology, the only drawback appears to be the time of electrodes setup (30 min instead of 10, which could be reduced in the future thanks to the use of dry electrodes and/or wireless EEG systems) and the increased data size (a less problematic issue nowadays). Simultaneous recordings of speech ABR and cortical AEP have been reported in only one study so far, but without any topographic EEG analysis (Musacchia, Strait, & Kraus, 2008). We agree with its authors on the fact that these two types of AEPs have different optimal stimulation parameters (number of presented trials and interstimulus interval). Although this might result in a drastically increased recording time to get both responses simultaneously, the issue can be bypassed by acquiring both markers in separate runs with respective optimal parameters. Furthermore, as shown in Figure 2e, multichannel recording of speech ABR allows using advanced processing methods like DSS, which allows for the

### Table 1. Comparison of Speech ABRs Obtained with the Two EEG Systems, and Benefits of Using DSS

<table>
<thead>
<tr>
<th>Number of accepted trials</th>
<th>Center</th>
<th>USB</th>
<th>BrainAmp</th>
<th>BrainAmp + DSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1,592</td>
<td>.41</td>
<td>12.2</td>
<td>12.8</td>
</tr>
<tr>
<td>S2</td>
<td>1,777</td>
<td>.56</td>
<td>5.4</td>
<td>8.0</td>
</tr>
<tr>
<td>S3</td>
<td>1,714</td>
<td>.61</td>
<td>5.8</td>
<td>1.2</td>
</tr>
<tr>
<td>S4</td>
<td>1,593</td>
<td>.33</td>
<td>4.2</td>
<td>-11.1</td>
</tr>
<tr>
<td>S5</td>
<td>1,650</td>
<td>.21</td>
<td>3.3</td>
<td>1.3</td>
</tr>
<tr>
<td>S6</td>
<td>1,648</td>
<td>.02</td>
<td>1.5</td>
<td>8.0</td>
</tr>
<tr>
<td>S7</td>
<td>1,335</td>
<td>.33</td>
<td>-5.7</td>
<td>-2.7</td>
</tr>
<tr>
<td>S8</td>
<td>1,710</td>
<td>.48</td>
<td>8.5</td>
<td>5.7</td>
</tr>
<tr>
<td>S9</td>
<td>1,719</td>
<td>.44</td>
<td>9.2</td>
<td>3.4</td>
</tr>
<tr>
<td>S10</td>
<td>1,611</td>
<td>.27</td>
<td>5.1</td>
<td>8.1</td>
</tr>
<tr>
<td>S11</td>
<td>1,775</td>
<td>.50</td>
<td>11.2</td>
<td>9.6</td>
</tr>
<tr>
<td>S12</td>
<td>1,264</td>
<td>.65</td>
<td>10.2</td>
<td>5.8</td>
</tr>
<tr>
<td>S13</td>
<td>2,483</td>
<td>.52</td>
<td>17.1</td>
<td>18.1</td>
</tr>
<tr>
<td>S14</td>
<td>1,531</td>
<td>.19</td>
<td>8.6</td>
<td>-4.5</td>
</tr>
<tr>
<td>Mean</td>
<td>1,672</td>
<td>.39</td>
<td>6.9</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Note. For each participant (S1 to S14), \(r^2\) is reported at a 0.75-ms lag (Center ahead of BrainAmp data), to take into account filter differences between the two systems (see text for details). SNR is reported for the classical vertical montage (Center and BrainAmp columns) and for the best DSS component (BrainAmp data only). SNR = signal-to-noise ratio; DSS = denoising source separation.
Topographic recordings of AEPs to speech

stimulus: /ba/

Speech ABR
vertex - mastoids

Speech ABR
best DSS component

Cortical AEP

SNR (dB)

research EEG system
vertical montage

research EEG system
best DSS component

clinical EEG system
vertical montage

Number of averaged trials

500 1000 all
reaching of high SNRs in only half the averaged trials required for classic processing. This is in agreement with a recent study using frequency-domain principal component analysis on multichannel recording of FFR to click trains, thus reducing by a factor of 3 the number of trials needed (Bharadwaj & Shinn-Cunningham, 2014).

Further studies are needed for assessing speech ABR generators using the present methodology. In the future, combining the use of speech ABR topographies and concomitant cortical AEPs should allow a better characterization of speech processing in populations with various hearing abilities. In any case, by proving the feasibility of recording high-quality speech ABR with standard EEG systems, and considering the widespread use of research EEG systems in neuroscience laboratories as well as the numerous fields of application of speech ABR, the present findings should amplify the development of speech ABR studies.

References


(Received March 11, 2014; Accepted September 7, 2014)