ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 41 (2008) 1313–1323

Evidence of the voice-related cortical potential: An electroencephalographic study

Jessica Galgano and Karen Froud*

Department of Biobehavioral Sciences, Teachers College, Columbia University, USA

Received 1 October 2007; revised 4 March 2008; accepted 12 March 2008
Available online 21 March 2008

The Bereitschaftspotential (BP) is a slow negative-going cortical potential associated with preparation for volitional movement. Studies since the 1960s have provided evidence for a BP preceding speech-related volitional motor acts. However, the BP associated specifically with voice initiation (i.e. a volitional motor act involving bilateral true vocal fold adduction) has not to date been systematically investigated. The current investigation utilizes a novel experimental design to address methodological confounds typically found in studies of movement-related cortical potentials, to demonstrate the existence and localization of generators for the voice-related cortical potential (VRCP). Using high-density EEG, we recorded scalp potentials in preparation for voice onset and for exhalation in a stimulus-induced voluntary movement task. Results showed a slow, increasingly negative cortical potential in the time window of up to 2500 ms prior to the mean onset of phonation. This VRCP peaked at a greater amplitude and shorter latency than the BP associated with exhalation alone. VRCP sources were localized to the anterior rostral regions of the medial frontal gyrus (Supplementary Motor Area (SMA)) and in bilateral laryngeal motor areas before and immediately following the mean initiation of phonation. Additional sources were localized to the bilateral cerebellum and occipital lobe in the time window following the mean onset of phonation. We speculate that these results provide additional support for fine somatotopic organization of the SMA. Further examination of the spatiotemporal change of the VRCP yielded source models which indicated involvement of the laryngeal motor cortices and cerebellum, likely responsible for the initiation and continuation of phonation.

Introduction

The event-preceding brain component associated with preparation for volitional movement, referred to as the Bereitschaftspotential (BP), has been described in detail over many years of research (Kornhuber and Deecke, 1965; Deecke et al., 1969, 1976).

Several studies have attempted to identify and isolate the BP related specifically to preparation for speech. For example, Brooker and Donald (1980) put a significant amount of consideration into matching the time constants of instrumentation, and included EMG recordings of several muscles that are active during speech. Wohlert (1993) and Wohlert and Larson (1991) investigated the BP preceding speech and nonspeech movements of various levels of complexity. Both experiments controlled for respiratory artifact by having subjects hold their breath prior to task initiation. In addition, electro-ocular and EMG activity were monitored, and (in the 1993 study) a pneumatic respiration transducer was utilized to monitor breathing patterns. Additionally, EMG activity from the orbicularis oris muscle was used to trigger and average segments. More recent advances in electroencephalographic and electromyographic techniques have made it possible for examinations of this nature to more accurately identify BPs associated with vocalization and oral movements.

These advances have also permitted investigations aiming to specify the cortical and subcortical pathways involved in volitional control of exhalation, which is required for voice production. Kuna et al. (1988) found thyroarytenoid muscle activity during exhalation, suggesting that cortical control of volitional respiration may be related, in part, to the requirement for precise management of vocal fold position during respiration.

Although a significant amount is understood about the BP, it has been difficult to extract these components from EEG recordings, since the BP is typically a slow change in amplitude with a wide bilateral distribution (Brooker and Donald, 1980; Deecke et al., 1986; Ertl and Schafer, 1967; Grabow and Elliott, 1974; McAdam and Whitaker, 1971; Morrell and Huntington, 1971; Schafer, 1967), representing shifts of only a few microvolts. Thus, accurate triggering by the exact onset of movement is extremely important. Studies attempting to identify the BP associated with the volitional motor act of laryngeal or vocal fold movement (which we will refer to as the Voice-Related Cortical Potential, or VRCP) have encountered other obstacles too: in particular, difficulties with co-

* Corresponding author. Department of Biobehavioral Sciences, Box 180, Teachers College, Columbia University, New York, NY 10027, USA. Fax: +1 212 678 8233.
E-mail address: kfroud@tc.columbia.edu (K. Froud).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.03.019

registration between physiological measurements and electrophysiological instrumentation, inaccurate identification of vocal fold movement onset, and methodological confounds between voice, speech and language (Brooker and Donald, 1980; Deecke et al., 1986; Ertl and Schafer, 1967; Grabow and Elliott, 1974; McAdam and Whitaker, 1971; Morrell and Huntington, 1971; Schafer, 1967). In addition, respiratory artifact or R-wave contamination of the BP preceding speech has proven a major difficulty, particularly in early studies (Deecke et al., 1986). Larger-amplitude artifacts due to head-, eye-, lip-, mouth movements and respiration must also be eliminated before signal averaging (Grözinger et al., 1980).

Earlier studies investigating voice-related brain activations typically confounded the distinctions between voice, speech and language. *Voice* refers to the sound produced by action of the vocal organs, in particular the larynx and its associated musculature. *Speech* is concerned with articulation, and the movement of organs responsible for the production of the sounds of language — in particular, those of the oral tract, including the lips, tongue and palate. *Language* refers to the complex set of cognitive operations involved in producing and understanding the systematic processes which underpin communication. Therefore, studies which have attempted to isolate voice or speech-related activity by the use of word production instead have described activation relating to a combination of these cognitive and motor operations (for example, Grözinger et al. (1975) used word utterances amongst their tasks designed to elicit speech-related activations; Ikeda and Shibasaki (1995) used single words as well as nonspeech-related movements like lingual protrusion; McAdam and Whitaker (1971) used unspecified three-syllable words to elicit ostensibly speech-related activity). Conversely, in a magnetoencephalography (MEG) study, Gunji et al. (2000) examined the vocalization-related cortical fields (VRCF) associated with repeated production of the vowel [u]. Microphones placed close to the mouth were used to capture the sound waveform from the vocalization; the onset of the waveform provided the trigger for segmenting and averaging epochs. This design carefully attempts to identify vocalization-related fields; however, operationalizing a procedure which is able to most closely capture the onset of voicing is particularly difficult. Difficulty stems, in part, from the limited number of compatible neuroimaging techniques and instruments able to capture these phenomena.

The present study contributes to understanding of the timing and distribution of the VRCP by addressing two major sources of methodological confound: the blurring of distinctions between voice, speech and language; and the accurate identification of movement onset for triggering and epoch segmentation. Furthermore, we use high-density EEG recordings, providing an increased level of detail in terms of the scalp topography, and additionally enabling the application of source modeling techniques to ensure accurate identification of the VRCP. Our results provide novel insight into voice generation by addressing the following research question:

> Can the true VRCP, associated only with laryngeal activity, be isolated from related movement potentials, by utilizing the right combination of control and experimental tasks?

We predicted that a stimulus-induced voluntary movement paradigm would yield significant differences in the characteristics of the Readiness Potentials associated with (a) initiation of phonation and (b) respiration. To be specific, we predicted the existence of an isolable voice-related cortical potential associated *only* with preparation for initiation of phonation and greater amplitude of the VRCP vs. the respiration-related cortical potential. We also predicted that VRCP sources would be localized to the Supplementary Motor Area, primary motor cortices, and sensori-motor regions.

Elucidation of the neural mechanisms of normal voice is a crucial step towards understanding the role of functional reorganization in cortical and subcortical networks associated with voice production, both for changes in the normal aging voice, and in pathological populations. This approach to determining the neural correlates of voice initiation could provide a foundation for creating neurophysiologic models of normal and disordered voice, ultimately informing our understanding of the effects of surgical, medicinal and/or behavioral interventions in voice-disordered populations. The findings could ultimately provide us with new basic science information regarding the relative benefit of different treatment approaches in the clinical management of neurogenic voice disorders. In addition, the larger significance of this work is related to the fact that voice disorders are currently recognized as the most common cause of communication difficulty across the lifespan, with a lifetime prevalence of almost 30% (Roy et al., 2005).

Materials and methods

A stimulus-induced voluntary movement paradigm in which trials of different types were presented in subject-specific randomized orders was utilized. This method addressed the documented problem of the classic BP paradigm which involves self-paced movements separated by short breaks: the person is already conscious of and preparing for a particular movement and there is a known repetition rate of the movements (Libet et al., 1982, 1983). This can lead to automatic movements, which change the presentation of the VRCP. In our procedure, it is not possible for the participant to predict ahead of time which task they have to perform, which allows for a spontaneous movement.

The movements were chosen to avoid another methodological confound, between voice, speech and language tasks. Requiring subjects to produce linguistically complex units, such as sounds or words (e.g. Ikeda and Shibasaki, 1995; Wohlert and Larson, 1991; Wohlert, 1993) led to some debate concerning whether BPs for speech might be lateralized to the dominant hemisphere for language. This problem is avoided in the current study, and the problem of movement artifacts involved in speech and speech-like movements such as lip-pursing or vowel-production (Gunji et al., 2000; Wohlert and Larson, 1991; Wohlert, 1993), in particular of back, tense, rounded vowels (such as the [u] used in Gunji et al's experiments), by utilizing a task which involves voicing *only*, and has no related speech or language overlay. The actions of breathing out through the nose, and gentle-onset humming of the bilabial nasal [m] without labial pressing, are equivalent actions in terms of involvement of the articulatory tract, the only difference being the initiation of vocal fold movement in the humming condition. By having participants breathe or hum following a period of breath-holding, the possibility of R-wave contamination is also reduced (Deecke et al., 1986). Onset of phonation is established by measuring vocal fold closure using electroglottography (EGG), and a telethermometer attached to a trans-nasal temperature probe was used for the earliest possible identification of exhalation onset.

Subjects

24 healthy subjects (21 females and 3 males) with an age range of 21–35 years of age (mean age=26 years) participated in the study. All subjects were informed of the purpose of the study and

gave informed consent to participate, following procedures approved by the local Institutional Review Board. All participants took part in a training phase, which was identical to the experimental procedure and served to train participants on the expected response to each screen. Each step of the procedure was discussed and explained as it was occurring, and there was ample opportunity for feedback to be provided to ensure accurate task performance.

EEG/ERP experimental set-up and procedures

EEG data acquisition

Scalp voltages were collected with a 128 channel Geodesic Sensor Net (Tucker, 1993) connected to a high-input impedance amplifier (Net Amps200, Electrical Geodesics Inc., Eugene, OR). Amplified analog voltages (.1–100 Hz bandpass) were digitized at 250 Hz. Individual sensors were adjusted until impedances were less than 30–50 kΩ, and all electrodes were referenced to the vertex (Cz) during recording. The net included channels above and below the eyes, and at the outer canthi, for identification of EOG. The EEG, EOG, stimulus triggered responses, EGG and telethermometer data were acquired simultaneously and later processed offline.

Recording of respiration

A nasal telethermometer (YSI Model 43 single-channel) with a small sensor (YSI Precision 4400 Series probe, style 4491A) was placed 2–4 cm inside one nostril transnasally and used to measure the temperature of inhaled and exhaled air. Readings from the telethermometer were digitally recorded by interfacing the telethermometer with one outrider channel input to the EEG net amplifier connection, for co-registration of the time course of respiration with the continuous EEG.

Recording of voice onset

A Kay Telemetric Computerized Speech Lab, Model 4500 (housing a Computerized Speech Lab Main Program Model 6103 Electroglottography) with 2 electrodes placed bilaterally on the thyroid cartilage, adjacent to the thyroid notch, was used to measure vocal fold closure and opening. The EGG trace was acquired in the Computerized Speech Lab (CSL) proprietary software and co registered offline with EEG and telethermometer recordings, in order to determine error trial locations and confirm onset of vocal fold adduction and controlled exhalation. Voice sounds were also recorded by microphone on a sound track acquired on the CSL computer, sampling at 44.1 kHz. A response button box permitted participant regulation of the start of each trial. At each button press, an audible "beep" was generated by the system which provided an additional point of co-registration between the EGG system and the time of trial onset. In addition, pressing the button permitted the subject to move to the next trial set from a screen that allowed physical adjustment into a more comfortable position if needed in between tasks (to reduce movement artifact).

Instructions and experimental task

The experimental task required subjects to hold their breath for 4 s, followed by breathing out or humming through the nose. The action carried out was determined by presentation of a "Go" screen after the breath-holding interval; the "Go" screen randomly presented either a "Breathe" or "Hum" instruction. To avoid using language-based stimuli in this experiment, the instructions to breath or hum were represented instead by letter symbols: a large 0 for breathing, and a large M for humming. There were eighty trials altogether (forty voice and forty breathe). Experimental stimuli were presented using Eprime stimulus presentation software (Psychology Software Tools, Pittsburgh, PA). Subjects were visually monitored via a closed circuit visual surveillance system, to ensure compliance with experimental conditions. Each trial (breathe or hum) was followed by a black screen, which indicated to participants that they could take a break before the next trial, swallow, blink and make themselves comfortable (this was intended to reduce movement artifacts during trials). Participants used button presses to indicate when they were ready to continue on to the next trial (Fig. 1).

Data analysis

Recorded EEG was digitally low-pass filtered at 30 Hz. Trials were discarded from analyses if they contained incorrect responses,

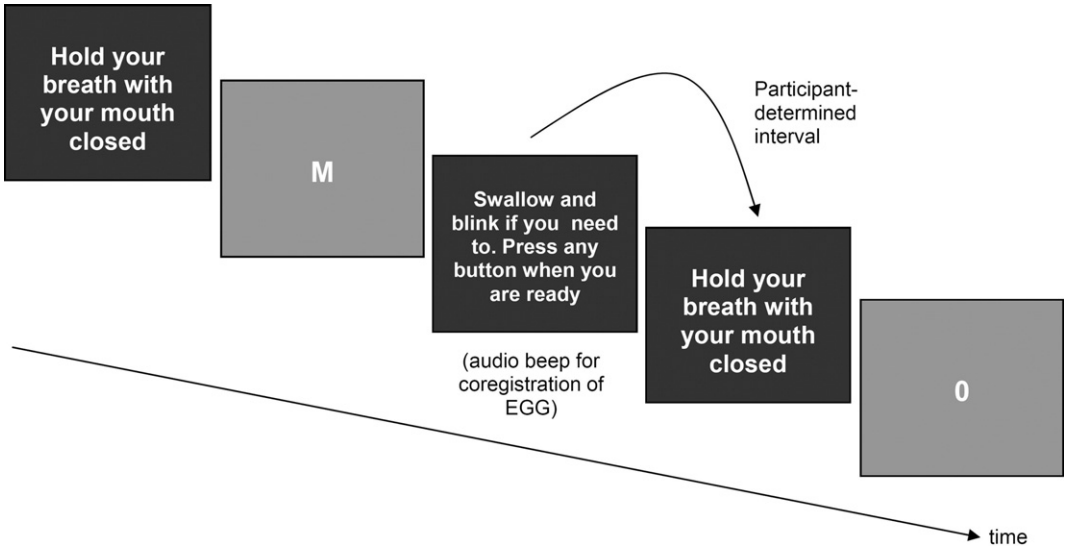


Fig. 1. The following experimental control module display shows the timeline of stimulus presentation during the experiment. Initially, a red screen instructed the subject to hold their breath with a closed mouth (4 s). This was followed by a green screen which displayed either an "M" or "0", prompting the subject to hum or breathe out, respectively. Following each trial, a black screen instructed the subjects to make themselves comfortable to minimize movement artifact before moving onto the next trial. When subjects were ready, a button press triggered an audio beep which allowed for co-registration of instrumentation being utilized.

eye movements (EOG over 70 μV), or more than 20% of the channels were bad (average amplitude over 100 μV). This resulted in rejection of less than 5% of trials for any individual. EEG was rereferenced offline to the average potential over the scalp (Picton et al., 2000). EEG epochs were segmented from −3000 to +500 ms from onset of voicing or exhalation, and averaged within subjects. Data were baseline-corrected to a 100 ms period from the start of the segment, to provide additional control for drift or other low amplitude artifact.

For identification of ERPs and for further statistical analyses, two regions of interest (ROIs) were selected: the Supplementary Motor Area (SMA) ROI, and the Primary Motor Region (M1) ROI. The 7 SMA sensors were centered around FCz, where SMA activations have previously been reported (e.g. Deecke et al., 1986). The M1 ROI consisted of 25 sensors, centered anterior to the central sulcus and located around the 10–20 system electrodes F7, F3, Fz, F4, F8, A1, T3, C3, Cz, C4, T4, A2 (listed left-to-right, anterior-to-posterior), where Motor-Related Potentials have been previously identified (Jahanshahi et al., 1995). See Fig. 2.

Statistical analyses

Data from averaged segments were exported to standard statistical software packages (Microsoft Excel and SPSS), permitting further analysis of the ERP data. Repeated measures Analysis of Variance (ANOVA) was used to evaluate interactions and main effects in a 2 (Condition: voicing vs. breathing)×2 (region: SMA vs. M1)×3 (time window: pre-stimulus, stimulus to voice onset, and post-voice onset) comparison. The dependent variable was grand-averaged voltages across relevant sensor arrays, determined following data preprocessing. The ANOVA was followed by planned comparisons, and all statistical analyses employed the Greenhouse–Geisser epsilon as needed to deal with violations of assumptions of sphericity.

Point-to-point differences in mean amplitude between the 2 conditions (humming vs. breathing) were evaluated for statistical significance, using separate repeated measures *t*-tests performed on mean amplitude measures within a 4 ms sliding analysis window. Bonferroni corrections were employed to control for type 1 error arising from multiple comparisons.

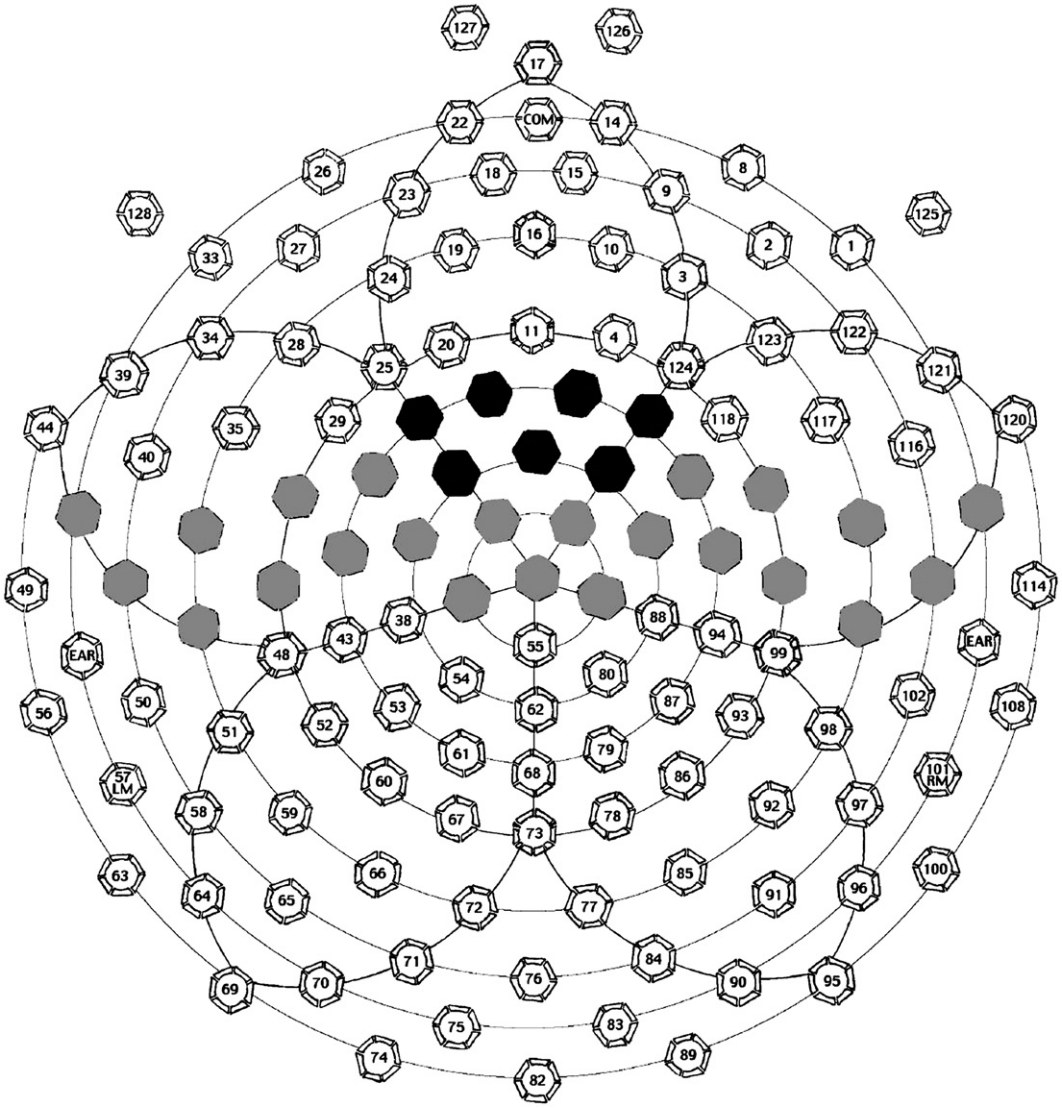


Fig. 2. This sensor layout displays the 128-channel Geodesic Sensor Net utilized in the current experiment. Legend: Black = SMA montage; Grey = M1 montage; Black + Grey = Channels entered into Grand Average.

VRCP. Time-locking of the segmented EEG to the onset of true vocal fold adduction as recorded from the electroglottograph enabled identification of the standard BP topography, with a peak at the time of the movement onset, followed by a positive reafferent potential.

The topography of the VRCP was examined using true vocal fold (TVF) adduction onset obtained from the EGG recording, and is subject-specific. Individual averaged files were placed into group grand-averages. The VRCP was identified in individual averaged data and in group grand-averages, based on the distribution and latency of activations. Component duration and mean amplitude for each subject (and for grand-averaged data) in each experimental condition were calculated.

Three pre-movement components of the VRCP were measured, i.e. early (−1500 to −1000 ms prior to movement onset), late (about −500 ms prior to movement onset), and peak VRCP (coincides with or occurs approximately 50 ms prior to movement onset) (Deecke et al., 1969, 1976, 1984; Barret et al., 1986). To determine the onset of each VRCP component, mean amplitude traces from individual and grand-averaged voice trials were examined independently by scientists with BP experience (Jahanshahi et al., 1995; Fuller et al., 1999). The mean latency of the early VRCP (rise of the slope from the baseline), the late VRCP (point of change in slope), and the peak VRCP (most negative point at or prior to vocal fold closure) were measured. The slope of the early VRCP was calculated (in microvolts per second) between the point of onset of the early VRCP and the onset of the late VRCP. The slope of the late component was calculated from the point of onset of the late VRCP to the onset of the peak VRCP. A 2 (region: SMA vs. M1) × 2 (time window: early VRCP vs. late VRCP/late VRCP vs. peak VRCP) repeated measures ANOVA, followed up with planned comparisons, was used to examine interactions and main effects.

BESA. In order to model the spatiotemporal properties of the VRCP sources, we used Brain Electrical Source Analysis (BESA: Scherg and Berg, 1991).

Source modeling procedures were applied to the voice production condition only (not to the exhalation condition). This is because a telethermometer was used to record changes in temperature associated with inhalation and exhalation; however, these associated changes do not reliably correlate with the true onset of exhalation or thyroarytenoid muscle activity associated with exhalation, as evidenced by the wide variety of measures reported in the literature for determination of respiration onset (e.g. Macefield and Gandevia (1991) used EMG measured over scalene and lateral abdominal muscles; Pause et al. (1999) used a thermistor placed at the nostril to determine onset of respiration based on changes to air temperature; Gross et al. (2003) determined onset of respiration to be associated with highest cyclic subglottal pressure; and other methods have also been reported). Source localization approaches are therefore not appropriate for the exhalation condition; consequently, we conducted comparisons between potentials associated with exhalation and voice using statistical analyses of differences in amplitude only. Source localization procedures were conducted on the voice production condition, because in that condition we were able to identify the initiation of voicing, using electroglottography.

BESA attempts to separate and image the principal components of the recorded waveform as well as localizing multiple equivalent current dipoles (ECDs). Any equivalent current dipole was fit to the data over a specified time window, and the goodness of fit was expressed as a percentage of the variance. Our procedure for developing the ECD model was closely based on procedures detailed in Gunji et al. (2000), as follows.

First, we selected an interval for analyzing the data in terms of a spatiotemporal dipole model. Following Gunji et al., we selected the interval of −150 ms to +100 ms, because this interval covered the approximate period from the onset of the instruction screen to preparation to move the vocal folds, through to onset of phonation and the start of auditory feedback. Gunji et al. further recommend a dipole modeling approach limited to this time interval in order to focus on brain activations just before and after vocalization, rather than attempting to model the complex and persistent sources associated with Readiness Potentials. We therefore seeded sources and fit them for orientation and location in the time window from −150 ms to 0 ms (the averaged time of the start of phonation). The time window from 0 to +100 ms was examined separately. Sources seeded in both time windows are described below.

Results

Individual data were grand-averaged and component identification was based on distribution, topography, and latency of activations (individual subjects and grand-averaged data). A VRCP was identified in all subjects, maximized over fronto-central electrodes (overlying the SMA). For grand-averaged data, all electrodes overlying the SMA showed a large VRCP in the specified time window (see Fig. 3).

Voicing vs. Controlled Exhalation Conditions

The ANOVA revealed that the triple Condition × Region × Time interaction was significant (F (1, 124) = 2488.463, $p < .0001$), as were both two-way interactions (Condition × Region, F (1, 124) = 68.428, $p < .0001$; Condition × Time, F (1, 124) = 1808.242, $p < .0001$; Region × Time, F (1, 124) = 6651.504, $p < .0001$). Planned comparisons revealed that the mean amplitudes of the VRCP were significantly more negative than the BP associated with the controlled exhalation condition, and SMA amplitudes were significantly more negative than M1. The significant interaction between Condition and Region for all subjects was found to be due to the fact that, although SMA sensors were always significantly more negative than M1 sensors (t (1939.233) = 26.272, $p < .0001$), there was a greater difference in the measured negativities in Voice trials compared to Breathe trials (see Fig. 4).

Further examination of the main effect of Time revealed that, as time progressed, mean amplitudes became significantly more negative (i.e. VRCPs became significantly more negative from the pre-stimulus time window to the time of voice onset and beyond). Investigations of the Condition by Time interaction revealed significant progressive increases in the measured negativities from early to late time windows for the Voice condition. However, subjects showed a greater degree of negativity in the pre- and post-screen time windows for the breathe condition only (see Fig. 5).

For the Controlled Exhalation/Breathing Condition, the SMA BPs from stimulus to exhalation were significantly more negative than in the pre-stimulus interval. The M1 region, however, showed no significant increase in negativity until the later time windows. In other words, over the SMA sensors the movement-related negativity increased in the period leading to exhalation, as well as later; over the M1 sensors, however, readings did not become significantly more negative until after movement. Investigations of the Region × Time interaction for the voicing trials showed a significant increase in the negativity over both the SMA and M1

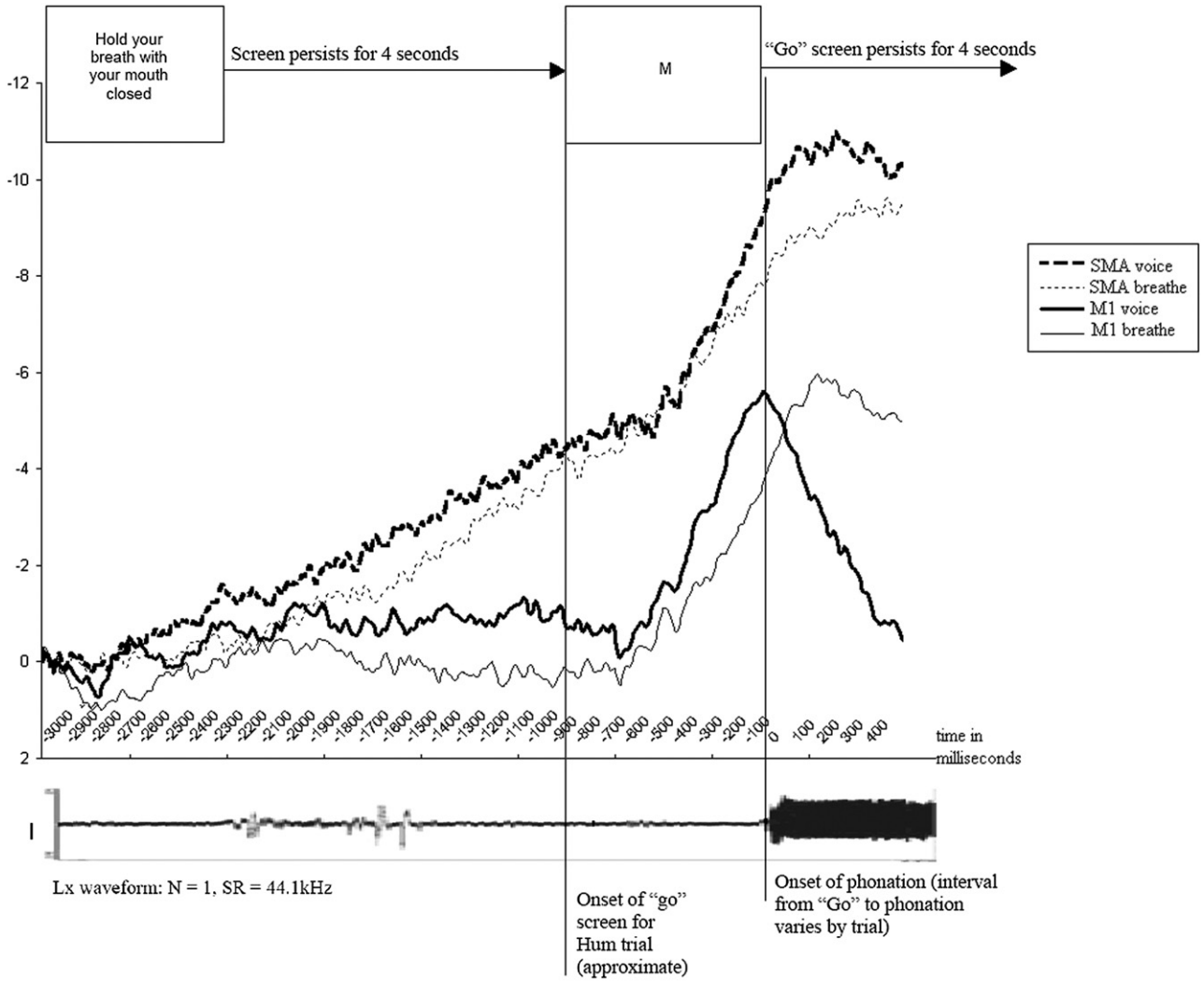


Fig. 3. The above waveform demonstrates grand-averages of 24 subjects. In the voice condition, a peak negativity of the VRCP (SMA: −10.0086, V; M1: −5.2983, V) was found at bilateral TVF adduction, evidenced by onset movement shown in the Lx (EGG) waveform. A standard BP topography in M1 is revealed. The late VRCP in M1 shows a steeper slope, positive deflection preceding movement onset, and longer latency when compared to SMA. In the breathe condition, peaks showed longer latencies over both M1 and SMA sensors, and reduced amplitude over SMA. Steps in the stimulus presentation/analysis procedure are superimposed: the breath-holding screen starts at −4000 ms, and the "Go" screen (instruction to hum through the nose) appears after 4 s of breath-holding and is shown for a further 4 s period. Initiation of phonation (recorded by EGG) was established for each individual trial within each subject. The interval between the onset of the "Go" screen and phonation is where the specific VRCP could be identified.

regions between the pre-stimulus interval and the time to voice onset. Mean amplitudes continued to become significantly more negative across time intervals post-voice onset for both regions. This is summarized in Table 1 and shown in Fig. 5 below.

To summarize, several significant findings were revealed. The voicing condition was significantly more negative than the exhalation condition, activation over SMA sensors was significantly more negative than over M1 sensors, and negativities significantly increased over the three time windows for the voice condition only.

VRCP slope changes

The 2 × 3 repeated measures ANOVA examining changes in the VRCP slope (microvolts per second) revealed a significant main effect of time, with the earlier time window being associated with a shallower slope than the later time window in both Regions. No other main effects or interactions were significant.

Source localization using BESA

Using BESA, we fit dipoles to the grand-averaged data from 23 subjects' responses to the Voice condition. We accepted an ECD model as a good fit when the residual variance dropped to 25% or below (standard for fitting to data from individuals is 10% RV). We began by seeding pairs of dipole sources to the left and right laryngeal motor areas, and in the middle frontal gyri, known to be associated with oro-facial movement planning in humans (Chainay et al., 2004) and the origination of human motor readiness potentials (Pedersen et al., 1998), respectively. A final pair of dipoles was

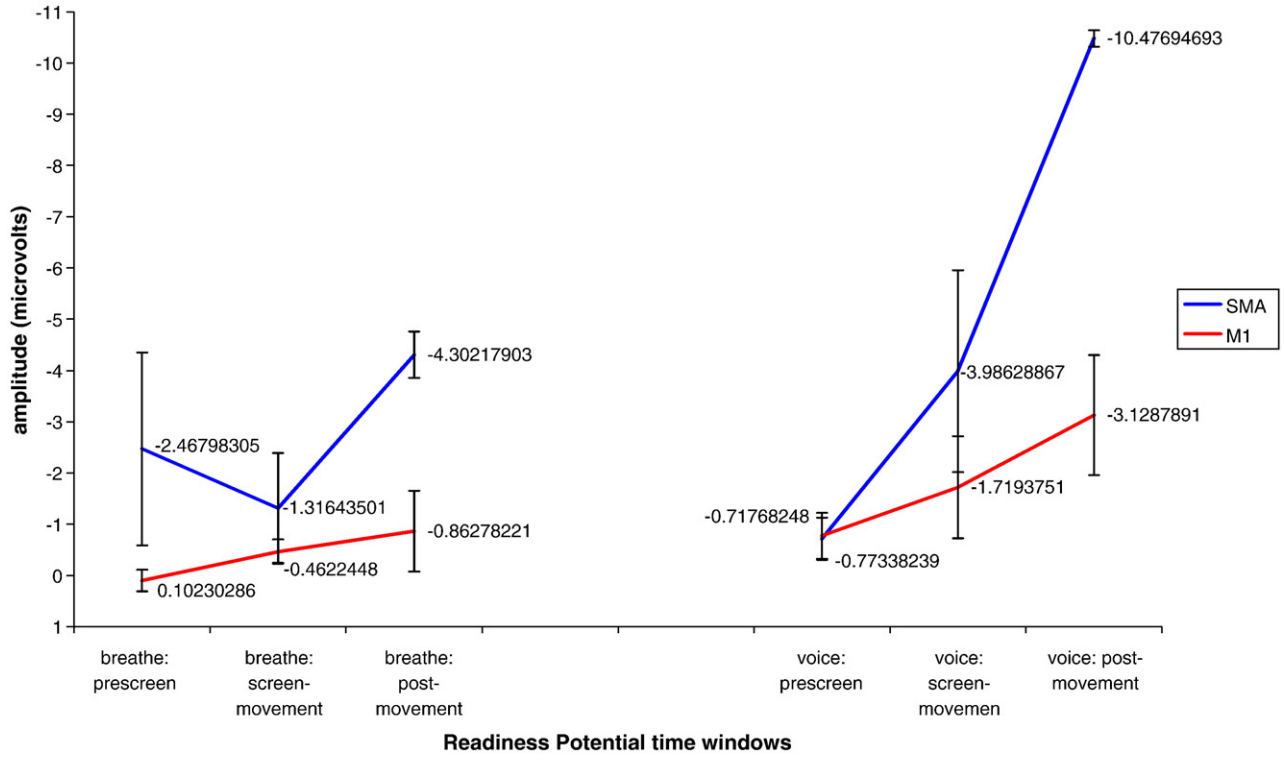


Fig. 4. This figure displays mean voltages over SMA and M1 sensors in three time windows. A significant reduction in negativity was shown over SMA sensors between RPs from the pre-stimulus interval to the stimulus to exhalation window for breathe trials only. A significantly greater degree of negativity was also seen in pre- and post-screen time windows in SMA for the breathe condition (prescreen: $n=194$, mean: −2.47 μV, SD: +/−1.88 μV; screen to movement: $n=432$, mean: −1.32 μV, SD: +/−1.07 μV; post-movement: $n=125$, mean: −4.30 μV, SD: +/−.45 μV). Apart from this, all differences between time windows represented significant increases in the negative-going potential. Significant progressive increases in the measured negativities from early to late time windows for the voice condition were evidenced in both regions (SMA: prescreen: $n=195$, mean: −.72 μV, SD: +/−.41 μV; screen to movement: $n=432$, mean: −3.99 μV, SD: +/−1.97 μV; post-movement: $n=125$, mean: −10.48 μV, SD: +/−.159 μV; M1: prescreen: $n=195$, mean: −.77 μV, SD: +/−.45 μV; screen to movement: $n=432$, mean: −1.72 μV, SD: +/−.99 μV; post-movement: $n=125$, mean: −3.13 μV, SD: +/−1.17 μV).

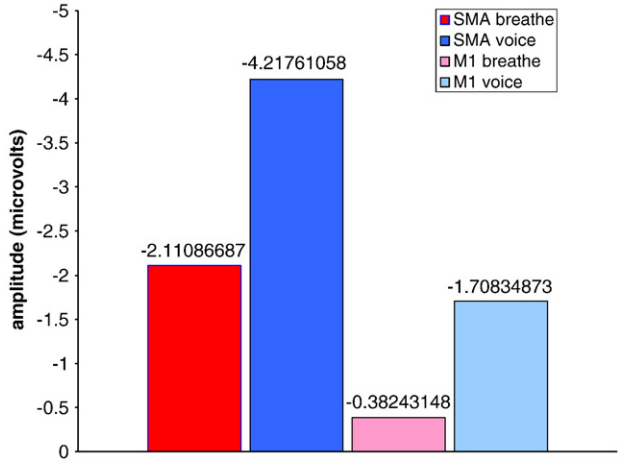


Fig. 5. This figure shows mean amplitudes over SMA and M1 sensors for all subjects. Greater degrees of negativity were observed in SMA and M1 regions for the voice condition. Greater degrees of negativity were also shown in SMA region for both conditions: Breathe/SMA: mean: −2.11 μV, SD: +/1.67 μV; Breathe/M1: mean: −.38 μV, SD: +/.50 μV; Voice/SMA: mean: −4.22 μV, SD: +/−3.47 μV; Voice/M1: mean: −1.71 μV, SD: +/−1.19 μV.

seeded in the occipital lobe. We modeled source orientation first and then performed independent fitting for location across the interval of −150 ms to 0 ms (Gunji et al., 2000). Residual variance with these six dipoles over this interval was 23.971% (Fig. 6).

Finally, we performed similar procedures for fitting equivalent current dipole sources over the interval from 0 (mean onset of phonation) to +100 ms (Gunji et al., 2000). In this interval, in addition to dipoles in the middle frontal gyri and primary motor areas, we seeded paired sources in the occipital lobe and cerebellum. Cerebellar sources have previously been found during vocalization and mouth movements (Milliken et al. (1999) and

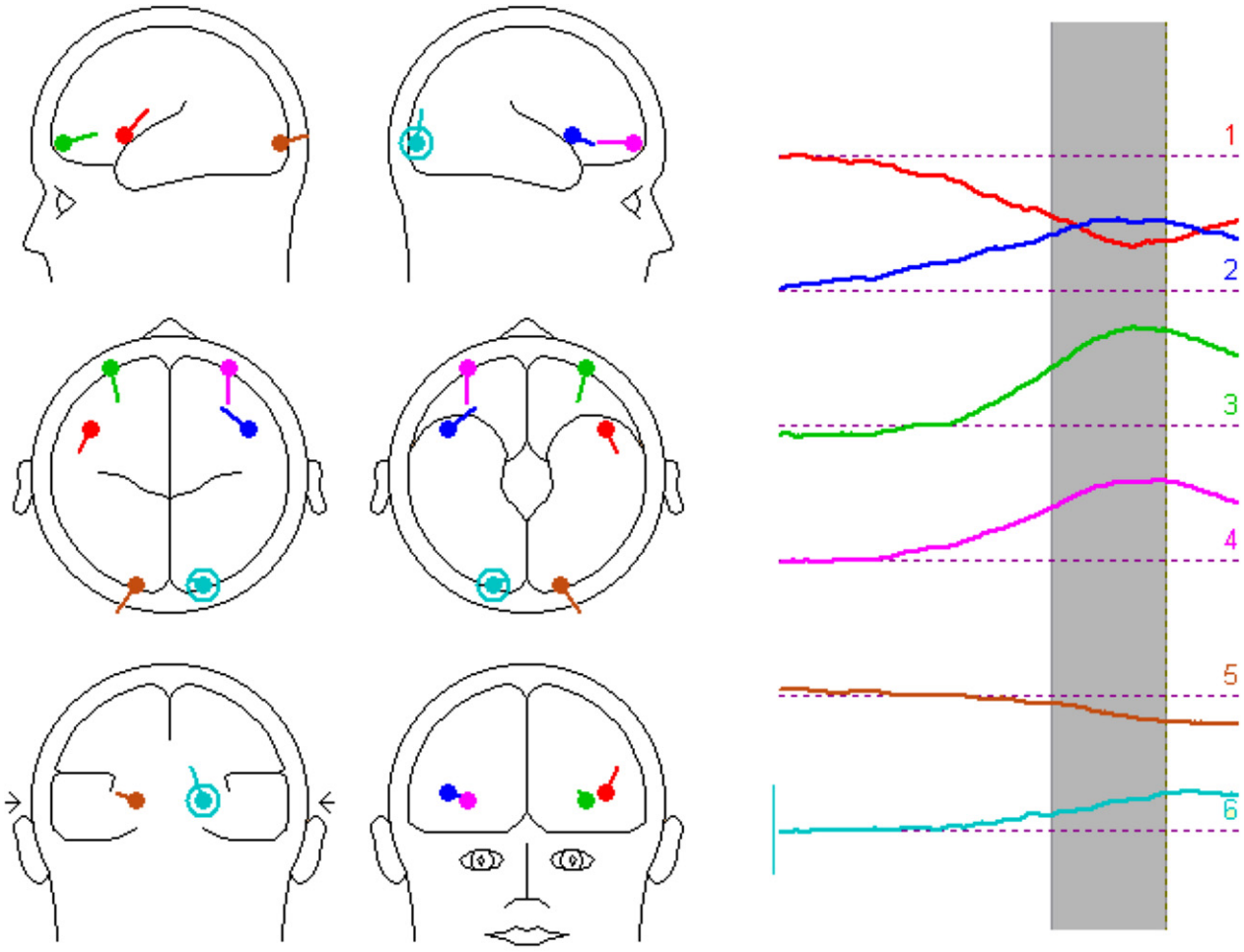
Table 1
Mean amplitudes for time windows across SMA and M1 sensors, for voice trials

Region/time period	Amplitude (μV)
SMA early	−1.8
SMA late	−4
SMA peak	−10
M1 early	−1
M1 late	−1.2
M1 peak	−5.3

Hirano et al., (1997)). These eight-dipole pairs accounted for 76.946% of the recorded variance in this time window (residual variance 23.054%) (Fig. 7).

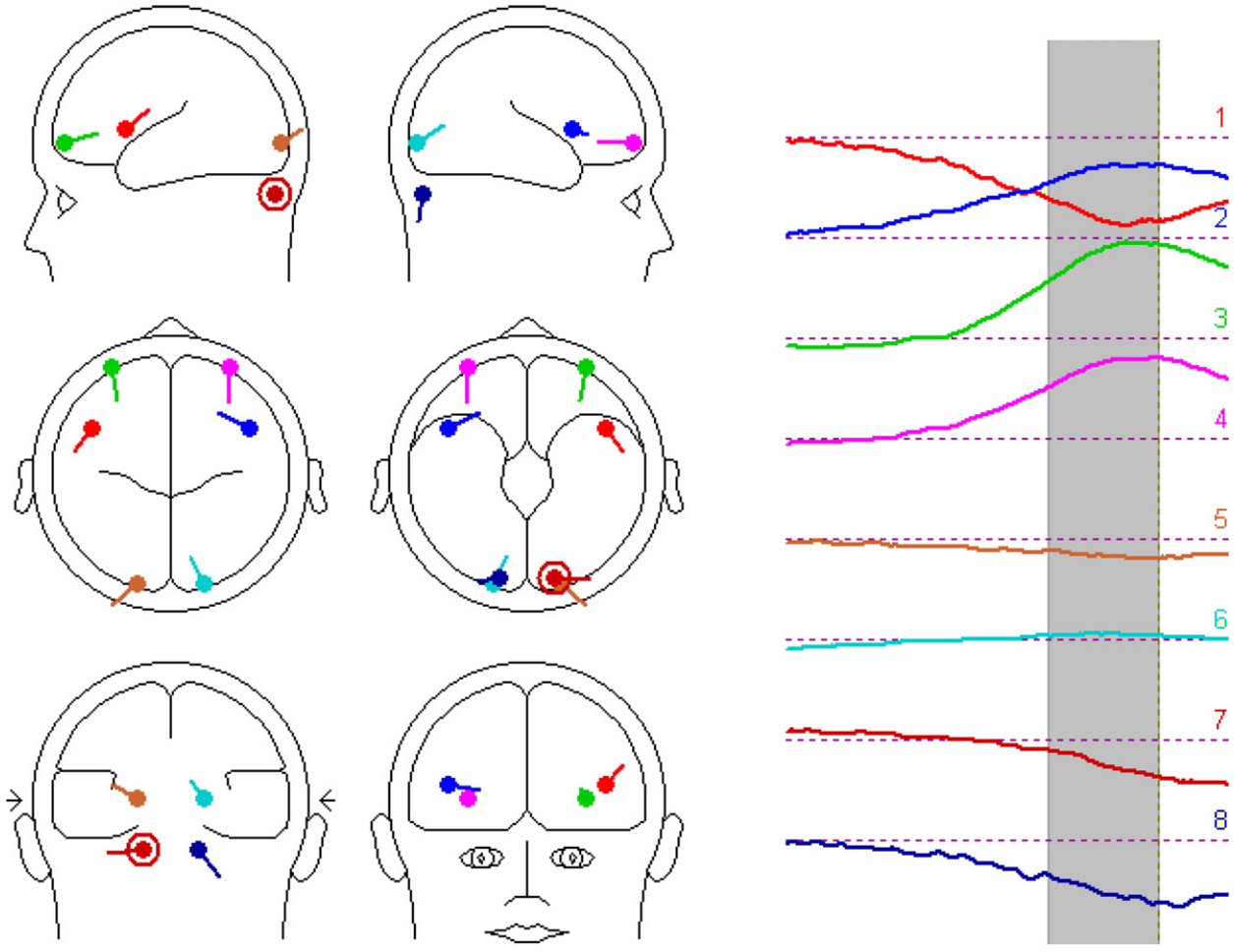
Discussion

To summarize, we have developed an experimental method which permits positive identification of the Voice-Related Cortical Potential (VRCP), while avoiding experimental confounds which have been problematic for previous investigations of this nature. A major point of the present study is that with the right combination of experimental and control tasks (breathing out through the nose vs. humming — identical motor tasks except for the phonation for /m/), the confound between speech, voice and language can be addressed and the VRCP associated only with voice production can be isolated. Leaving the instruction to hum or breathe until the last minute meant that participants were not able to consciously prepare for one action or the other. The BP for breathing or humming was therefore specific to each condition only in the few hundred milliseconds between the "go" screen presentation, and



Source	x-location	y-location	z-location	Region
1	-48.40	16.32	9.00	Left precentral gyrus
2	48.40	16.32	9.00	Right precentral gyrus
3	-36.80	58.30	8.90	Left middle frontal gyrus
4	36.80	58.30	8.90	Right middle frontal gyrus
5	-21.10	-96.10	-8.40	Left middle occipital gyrus
6	21.10	-96.10	-8.40	Right middle occipital gyrus

Fig. 6. This figure displays source modeling results for the time window 0 ms to +100 ms (grand averaged data). This six-dipole model yielded 23.97% residual variance over the pre-motion time window. Localization and orientation of each source on the spherical head model, as estimated by BESA (top right), is shown. The line from each point indicates the current direction for each dipole. The current waveform at each source is shown at left. The table shows source localizations and regions determined via the Talairach Daemon client. The *x*-axis indicates the anterior-posterior direction, with positive *x* coming out of the head at the nasion. The *z*-axis indicates the ventral direction, with positive values toward the upper side. The *y*-axis indicates the vertical direction with positive values toward the left pre-auricular point.



Source	x-location	y-location	z-location	Region
1	-48.40	16.32	9.00	Left precentral gyrus
2	48.40	16.32	9.00	Right precentral gyrus
3	-36.80	58.30	8.90	Left middle frontal gyrus
4	36.80	58.30	8.90	Right middle frontal gyrus
5	-21.10	-96.10	-8.40	Left middle occipital gyrus
6	21.10	-96.10	-8.40	Right middle occipital gyrus
7	-21.76	-74.42	-40.41	Left cerebellum
8	21.76	-74.42	-40.41	Right cerebellum

Fig. 7. This figure displays source modeling for time window 0 ms to +100 ms (grand averaged data). This eight-dipole model yielded 23.05% residual variance over the post-motion time window. Localization and orientation of each source on the spherical head model, as estimated by BESA (top right), is shown. The line from each point indicates the current direction for each dipole. The current waveform at each source is shown at left. The table shows source localizations and regions determined via the Talairach Daemon client. The *x*-axis indicates the anterior-posterior direction, with positive *x* coming out of the head at the nasion. The *z*-axis indicates the ventral direction, with positive values toward the upper side. The *y*-axis indicates the vertical direction with positive values toward the left pre-auricular point.

the defined onset of exhalation or phonation. Giving any earlier notice of the required action would have more likely resulted in isolation of a contingent negative variation (CNV) — a component known to be associated with anticipation of the upcoming stimulus and planning to respond (Ikeda et al., 1997).

Our experimental procedure also allowed for minimization of R-wave and predictability artifacts which are known to compromise the presentation of the BP in investigations of voice-related cortical activity. Furthermore, utilization of an electroglottogram (EGG) permitted accurate identification of vocal fold movement specific to voicing which in turn enabled accurate segmentation of epochs which were time-locked to the onset of phonation.

By making use of high-density EEG recordings, source modeling procedures were carried out which provided a further level of confirmation that the observed component was indeed a BP, involving generators in the SMA, middle frontal gyri and in M1, bilaterally. As

Shibasaki and Hallett (2006) point out, the BP is typically recorded under self-paced conditions, so it is unclear whether movements with more naturalistic and spontaneous presentations are associated with BP or not. However, our experiment did not use a self-paced paradigm, rather relying on voice produced in response to an external, and non-predictable, trigger. This work therefore lays groundwork for examination of the relationships between neural, physiological and psychological mechanisms involved in voice, movement and intentionality.

Examination of the spatiotemporal change of the VRCP using BESA yielded a six-dipole model prior to the onset of phonation (i.e. time period from −150 ms to 0 ms) and an eight-dipole model following the initiation of phonation (i.e. time period from 0 ms to 100 ms).

The sources found in the middle frontal gyri are consistent with studies indicating activity within these regions during the preparation period for human motor planning (e.g. Pedersen et al., 1998). Specifically, recruitment of these fronto-orbital regions is thought to reflect the cognitive action of choosing a particular movement to perform (Deecke and Kornhuber, 1978). Since the present study involves a stimulus-induced voluntary movement paradigm, this activation could be related to recognition and decision factors following on from stimulus presentation.

Dipole source analysis applied to the time window of −150 ms to 0 ms did not indicate sources in the SMA. Other investigators who have examined the BP have presented dipole models that did not include SMA sources (Bocker et al., 1994; Bötzel et al., 1993). However, given the strong SMA region activity found in the present study's region analysis investigating the Region × Time interactions for voicing trials, it seems unlikely that there are no source dipoles in the SMA. Examinations utilizing intracranial recordings have also reported SMA activity prior to volitional movement, providing further support of SMA contribution to generation of the BP (Ikeda et al., 1992; Rektor et al., 1994, 2001). Therefore, it seems plausible that the absence of SMA dipoles in our source model may be due to insufficient separation, in the current investigation, between SMA and M1 activity (see also Erdler et al., 2000).

The source model indicated involvement of the laryngeal motor cortices, which are likely responsible for the execution and continuation of phonation. These sources were localized to the laryngeal areas of the 'homunculus' of the sensori-motor cortex (Penfield and Boldrey, 1937).

Sources were also localized to the cerebellum, and (in the post-phonation time window) the occipital lobe bilaterally (in both time windows). This is consistent with results from Milliken et al. (1999) and Hirano et al. (1997), who reported activations in visual cortex, Broca's area, supplementary motor area (SMA), left Heschl's gyrus and the cerebellum during the vocalization of many kinds of sentences and during mouth movements, respectively.

All of our sources were bilaterally positioned pairs of dipoles. Sensory and motor oro-facial functions are known to have bilateral representation. Some investigators have reported left-lateralized activations (e.g. of Broca's area) during syllable or word production (Grözinger et al., 1980). However, the experimental task used in the present study involved a clear contrast between two physical rather than linguistic activities, equivalent in almost every way except that one (voicing) involved laryngeal movement. Therefore, we could be reasonably sure that we were isolating activations associated only with vocalization rather than aspects of speech or language production.

The current dipole source model found no additional generators that accounted for further variance. This suggests that the model presented here is exhaustive in terms of identifying the major contributing sources of the VRCP. However, we could not categorically exclude activity in other regions.

It is interesting to note that the slope of the late VRCP in M1 is much steeper than that in SMA (see Fig. 3). This may reflect the complexity of motor movements and control necessary for voicing (Shibasaki and Hallett, 2006). Furthermore, the positive deflection of the late VRCP in M1 may evidence inhibition of intrinsic and extrinsic laryngeal musculature required to initiate and maintain voicing at a steady frequency/pitch. In addition, the offset of the late VRCP in M1 (−400 ms) is later than that in SMA (−600 ms), possibly reflecting the later engagement of the primary motor areas following motor preparation in the premotor areas.

In conclusion, we developed a novel experimental design to address methodological confounds typically found in studies of movement-related cortical potentials, to demonstrate the existence and localization of generators for the voice-related cortical potential (VRCP). Using high-density EEG, we recorded scalp potentials in preparation for voice onset and for exhalation in a stimulus-induced voluntary movement task. Results showed a slow, increasingly negative cortical potential in the time window preceding the onset of phonation. This VRCP peaked at a greater amplitude and a shorter latency than the BP associated with exhalation alone. VRCP sources were successfully localized to the middle frontal gyri and in bilateral laryngeal motor areas before and immediately following the mean initiation of phonation. Additional sources were localized to the left and right cerebellum and occipital lobe in the 100 ms time window following the mean onset of phonation. Further examination of the spatiotemporal change of the VRCP yielded source models indicating involvement of the laryngeal motor cortices and cerebellum, likely responsible for the initiation and continuation of phonation.

Acknowledgments

We would like to acknowledge Benjamin Balas' and Kachina Allen's work on Matlab scripting. Thanks also to Mark Moran at EGI, Ronald DeMeersman and Greg Gates for technical assistance; and to all in the Neurocognition of Language Lab for support and many valuable discussions. In addition, we are very grateful to two anonymous reviewers for detailed and insightful comments on earlier drafts.

References

Barret, G., Shibasaki, H., Neshige, R., 1986. Cortical potentials preceding voluntary movement: evidence for three periods of preparation in man. Electroencephalogr. Clin. Neurophysiol. 63, 327–339.

Bocker, H., Kleinschmidt, A., Requardt, M., Hanicker, W., Merboldt, K.D., Frahm, J., 1994. Functional cooperativity of human cortical motor areas during self-paced simple finger movements. A high-resolution MRI study. Brain 117, 1231–1239.

Bötzel, K., Plendl, H., Paulus, W., Scherg, M., 1993. Bereitschaftspotential: is there a contribution of the supplementary motor area? Electroencephalogr. Clin. Neurophysiol. 89, 187–196.

Brooker, B.H., Donald, M.W., 1980. Contribution of the speech musculature to apparent human EEG asymmetries prior to vocalization. Brain Lang. 9 (2), 226–245.

Chainay, H., Krainik, A., Tanguy, M., Gerardin, E., Le Bihn, D., Lehericy, S., 2004. Foot, face and hand representation in the human supplementary motor area. Neuroreport 15, 765–769.

Deecke, L., Kornhuber, H.H., 1978. An electrical sign of participation of the mesial 'supplementary' motor cortex in human voluntary finger movements. Brain Res. 159, 473–476.

Deecke, L., Scheid, P., Kornhuber, H.H., 1969. Distribution of readiness potential, promotion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements. Exp. Brain Res. 7, 158–168.

Deecke, L., Grözinger, B., Kornhuber, H.H., 1976. Voluntary finger movement in man: cerebral potentials and theory. Biol. Cybern. 23, 99–119.

Deecke, L., Bashore, T., Brunia, C.H., Grunewald-Zuberbier, E., Grunewald, G., Kristeva, R., 1984. Movement-associated potentials and motor control. Report of the EPIC VI Motor Panel. Ann. N.Y. Acad. Sci. 425, 398–428.

Deecke, L., Engel, M., Lang, W., Kornhuber, H.H., 1986. Bereitschaftspotential preceding speech after holding breath. Exp. Brain Res. 65 (1), 219–223.

Erdler, M., Beisteiner, R., Mayer, D., Kaindl, T., Edward, V., Windschberger, C., Lindinger, G., Deecke, L., 2000. Supplementary motor area activation preceding voluntary movement is detectable with a whole-scalp magnetoencephalography system. Neuroimage 11, 697–707.

Ertl, J., Schafer, E.W., 1967. Cortical activity preceding speech. Life Sci. 6 (5), 473–479.

Fuller, R., Jahanshahi, M., Nathaniel-James, D., 1999. Movement-related potentials prior to self-initiated movements are impaired in patients with schizophrenia and negative signs. Exp. Brain Res 126 (4,), 545–555.

Grabow, J.D., Elliott, F.W., 1974. The electrophysiologic assessment of hemispheric asymmetries during speech. J. Speech Hear. Res. 17 (1), 64–72.

Gross, R.D., Atwood Jr., C.W., Grayhack, J.P., Shaiman, S., 2003. Lung volume effects on pharyngeal swallowing physiology. J. Appl. Physiol. 95 (6), 2211–2217.

Grözinger, B., Kornhuber, H.H., Kriebel, J., 1975. Methodological problems in the investigation of cerebral potentials preceding speech: determining the onset and suppressing artefacts caused by speech. Neuropsychologia 13 (3), 263–270.

Grözinger, B., Kornhuber, H.H., Kriebel, J., Szirtes, J., Westphal, K.T., 1980. The Bereitschaftspotential preceding the act of speaking. Also an analysis of artifacts. Prog. Brain Res. 54, 798–804.

Gunji, A., Kakigi, R., Hoshiyama, M., 2000. Spatiotemporal source analysis of vocalization-associated magnetic fields. Cogn. Brain Res 9 (2), 157–163.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., Konishi, J., 1997. Cortical processing mechanism for vocalization with auditory verbal feedback. NeuroReport 8, 2379–2382.

Ikeda, A., Shibasaki, H., 1995. Noninvasive study of cortical neuronal mechanism in voluntary movements: role of basal ganglia and cerebellum. J. Clin. Neurol. 35 (12), 1522–1524.

Ikeda, A., Lüders, H.O., Burgess, R.C., Shibasaki, H., 1992. Movement-related potentials recorded from supplementary motor area and primary motor area. Brain 115, 1017–1043.

Ikeda, A., Shibasaki, H., Kaji, R., Terada, K., Nagamine, T., Honda, M., et al., 1997. Dissociation between contingent negative variation (CNV) and Bereitschaftspotential (BP) in patients with parkinsonism. Electroencephalogr. Clin. Neurophysiol. 102 (2), 142–151.

Jahanshahi, M., Brown, R.G., Passingham, R., Brooks, D., Frackowiak, R.S.J., Marsden, C.D., 1995. Self-initiated versus externally triggered movements. I. An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. Brain 118 (4), 913–933.

Kornhuber, H.H., Deecke, L., 1965. Changes in the brain potential in voluntary movements and passive movements in man: readiness potential and reafferent potentials. Pflugers Arch. Gesamte Physiol. Menschen Tiere 284, 1–17.

Kuna, S.T., Insalaco, G., Woodson, G.E., 1988. Thyroarytenoid muscle activity during wakefulness and sleep in normal adults. J. Appl. Physiol. 65 (3), 1332–1339.

Libet, B., Wright Jr., E.W., Gleason, C.A., 1982. Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts. Electroencephalogr. Clin. Neurophysiol. 54 (3), 322–335.

Libet, B., Gleason, C.A., Wright, E.W., Pearl, D.K., 1983. Time of conscious intention to act in relation to onset of cerebral activity. Brain 102, 623–642.

McAdam, D., Whitaker, H.A., 1971. Language production: electroencephalographic localization in the normal human brain. Science 172, 499–502.

Macefield, G., Gandevia, S.C., 1991. The cortical drive to human respiratory muscles in the awake state assessed by premotor cerebral potentials. J. Physiol. 439, 545–558.

Milliken, G.W., Dobrivoje, S., Tarkka, I.M., 1999. Sources of movement-related cortical potentials derived from foot, finger, and mouth movements. Nonconvulsive status epilepticus. J. Clin. Neurophysiol. Nonconvulsive Status Epilepticus 16 (4), 361–372.

Morrell, L., Huntington, D., 1971. Electrocortical localization of language production. Science 174, 1359–1360.

Pause, B.M., Krauel, K., Sojka, B., Ferstl, R., 1999. Is odor processing related to oral breathing? Int. J. Psychophysiol 32 (3), 251–260.

Pedersen, J.R., Johanssen, P., Bak, C.K., Kofoed, B., Saermark, K., Gjedde, A., 1998. Origin of human motor readiness field linked to left middle frontal gyrus by MEG and PET. NeuroImage 8, 214–220.

Penfield, W., Boldrey, E., 1937. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain 60, 389–443.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson Jr., R., et al., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37 (2), 127–152.

Rektor, I., Feve, A., Buser, P., Bathien, N., Lamarche, M., 1994. Intracerebral recording of movement related readiness potentials: an exploration in epileptic patients. Electroencephalogr. Clin. Neurophysiol. 90 (4), 273–283.

Rektor, I., Bares, M., Kanovsky, P., Kukleta, M., 2001. Intracerebral recording of readiness potential induced by a complex motor task. Mov. Disord. 16 (4), 698–704.

Roy, N., Merrill, R.N., Gray, S.D., Smith, E.M., 2005. Voice disorders in the general population: prevalence, risk factors, and occupational impact. Laryngoscope 115 (11), 1988–1995.

Schafer, E.W., 1967. Cortical activity preceding speech: semantic specificity. Nature 216 (122), 1338–1339.

Scherg, M., Berg, P., 1991. Use of prior knowledge in brain electromagnetic source analysis. Brain Topogr. 4 (2), 143–150.

Shibasaki, H., Hallett, M., 2006. What is the Bereitschaftspotential? Clin. Neurophysiol. 17, 2351–2356.

Tucker, D.M., 1993. Spatial sampling of head electrical fields: the geodesic sensor net. Electroencephalogr. Clin. Neurophysiol. 87 (3), 154–163.

Wohlert, A.B., 1993. Event-related brain potentials preceding speech and nonspeech oral movements of varying complexity. J. Speech Hear. Res. 36 (5), 897–905.

Wohlert, A.B., Larson, C.R., 1991. Cerebral averaged potentials preceding oral movement. J. Speech Hear. Res. 34 (6), 1387–1396.