CrossMark

ORIGINAL ARTICLE

Spatiotemporal differentiation in auditory and motor regions during auditory phoneme discrimination

Annelies Aerts[1,2] · Gregor Strobbe[3] · Pieter van Mierlo[3] · Robert J. Hartsuiker[4] ·
Paul Corthals[5,6] · Patrick Santens[1,2] · Miet De Letter[2,5]

Abstract Auditory phoneme discrimination (APD) is supported by both auditory and motor regions through a sensorimotor interface embedded in a fronto-temporo-parietal cortical network. However, the specific spatiotemporal organization of this network during APD with respect to different types of phonemic contrasts is still unclear. Here, we use source reconstruction, applied to event-related potentials in a group of 47 participants, to uncover a potential spatiotemporal differentiation in these brain regions during a passive and active APD task with respect to place of articulation (PoA), voicing and manner of articulation (MoA). Results demonstrate that in an early stage (50–110 ms), auditory, motor and sensorimotor regions elicit more activation during the passive and active APD task with MoA and active APD task with voicing compared to PoA. In a later stage (130–175 ms), the same auditory and motor regions elicit more activation during the APD task with PoA compared to MoA and voicing, yet only in the active condition, implying important timing differences. Degree of attention influences a frontal network during the APD task with PoA, whereas auditory regions are more affected during the APD task with MoA and voicing. Based on these findings, it can be carefully suggested that APD is supported by the integration of early activation of auditory-acoustic properties in superior temporal regions, more perpetuated for MoA and voicing, and later auditory-to-motor integration in sensorimotor areas, more perpetuated for PoA.

Keywords Mismatch negativity (MMN) · P300 · Neurophysiology · Phonology

Introduction

Auditory phoneme identification, segmentation and discrimination during speech processing requires explicit access to certain sublexical speech segments, in contrast to speech processing in the context of more holistic speech recognition and comprehension, and is supported by a dorsal processing stream running from posterior superior temporal gyrus (STG) via inferior parietal regions towards frontal regions [1–3]. Initial, primary acoustic–phonetic analysis is supported by the middle and posterior part of the STG and superior temporal sulcus (STS) bilaterally, although hemispheric asymmetries do occur [1, 2, 4–8]. Left superior temporal areas ought to be specialized for processing temporal changes, whereas the right homologue areas should focus on spectral analyses [9]. Continuing with the time domain, the left hemisphere processes shorter timescales (25–50 ms; e.g. extraction of segment boundaries) and the right hemisphere longer timescales (200–300 ms; e.g. sentential prosody) [10]. From then on,

✉ Annelies Aerts
  aaerts@me.com

1 Department of Internal Medicine, Ghent University, De Pintelaan 185, 9000 Ghent, Belgium

2 Department of Neurology, Ghent University Hospital, Ghent, Belgium

3 Department of Electronics and Information Systems, Medical Image and Signal Processing Group, Ghent University-IMinds, Ghent, Belgium

4 Department of Experimental Psychology, Ghent University, Ghent, Belgium

5 Department of Speech, Language and Hearing Sciences, Ghent University, Ghent, Belgium

6 Department of Education, Health and Social Work, University College Ghent, Ghent, Belgium

activation spreads from the posterior STG via the inferior parietal cortex to the motor cortex and inferior frontal gyrus (BA 44/45) along the superior longitudinal and arcuate fascicle [2, 3, 7, 11, 12]. Within the dorsal pathway, the inferior parietal cortex serves as an auditory-motor interface between auditory features and articulatory-based representations in achieving successful detection of phonological changes [13–15]. Several studies have shown that the premotor and primary motor regions, usually active during speech production, are differentially activated in a somatotopic manner, based on articulatory characteristics of the phonemic contrasts, during phoneme perception and discrimination processes (e.g. activation of lip area during perception of bilabial place of articulation) [16–20]. Moreover, inferior frontal regions have been implicated in speech recognition as well [21, 22], when a new, non-familiar word or a pseudoword is encountered. Such word stimuli require more intensive segmentation processes and articulatory-based representations of the different segments (i.e. phonemes), hence sublexical processing, which involves more participation of frontal and motor areas within the dorsal circuitry [11, 23–25].

Importantly, by measuring event-related potentials (ERPs) and estimating the source of this event-related electrical activity through non-invasive source reconstruction, some interesting findings have been revealed regarding the spatiotemporal neural activity during phoneme perception. Already 100 ms after phoneme presentation, auditory association cortices (STG) are supposed to be spatially mapped along an anterior-posterior dimension determined by mutually exclusive place of articulation (PoA) features, as derived from the N100m (magnetic counterpart of the N100) and Mismatch Field (magnetic counterpart of the Mismatch Negativity) [26–29]. Moreover, even earlier, around 50 ms after phoneme presentation, neuronal generators of the P50 m are topographically aligned in more primary auditory regions, such as Heschl's gyrus, along a medial–lateral dimension according to PoA features [30]. Generally, ERPs like the P50 and N100 potential represent earlier, sensory-perceptual processes of phoneme detection and identification [30, 31], while ERPs like the (pre-attentive) Mismatch Negativity (MMN; elicited during passive oddball paradigm) and (attentive) P300 potential (elicited during active oddball paradigm) are more associated with phoneme discrimination processes [32, 33]. However, by comparing the neuronal generators of the MMN in the conventional design (distraction by video) and more attention-demanding condition (distraction by video, yet response to target requested), increased attention to the auditory stimuli was linked to greater participation of motor areas and inferior frontal regions during later stages (200–400 ms) of the discrimination processes [34].

Clearly, phonemic contrasts play a significant role in terms of neural activity in auditory and motor regions during phoneme perception and discrimination. Especially for PoA an important interaction has been demonstrated between the spatial mapping and temporal organization of neural activity associated with PoA features [27, 30]. For instance, source reconstruction of the N100m showed that stop-consonants with a labial PoA are processed more anterior on the superior temporal horizontal axis, whereas stop-consonants with a dorsal PoA are processed more posterior on the superior temporal horizontal axis [27]. However, it is remarkable that phonemes have only been compared at opposite sides of one phonemic contrast continuum (e.g. /d/ and /g/ for PoA) with respect to their somatotopic organization in auditory and motor cortices, whereby the focus mainly was on the PoA features [16, 30]. This can be related to the more unambiguous, invariable articulatory gestures inherent to PoA features for which potentially more clear-cut spatial representations exist in the brain [35]. Such motor properties are less well distinguishable for phonemic contrasts like voicing (e.g. /d/ and /t/) and manner of articulation (MoA) (e.g. /d/ and /z/), for which perception and discrimination processes rely more on well-defined acoustic cues [36]. On the contrary, acoustic cues of the PoA contrast are highly variable depending on the phonetic context, whereby invariant acoustic features are missing.

The present study aims to investigate on the one hand whether phoneme discrimination based on PoA elicits more motor cortical activity, considering its limited invariant acoustic cues and clear-cut motor representations, and on the other hand whether phoneme discrimination based on voicing or MoA evokes more auditory cortical activity, considering their well-defined acoustic features and unclear motor representations. Because of the important temporal aspect of neural activity during speech processing [30, 37, 38], the current study implements neurophysiological measures (MMN, P300) in response to an auditory oddball paradigm on which source reconstruction is performed. As such, it is examined whether a spatial cortical pattern associated with a particular phonemic contrast occurs with a certain temporal differentiation compared to the other phonemic contrasts. A second part of the study was to inspect whether increased attention can be related to higher activation of frontal areas [34], when the passively elicited MMN (conventional design) and actively elicited P300 are compared. These hypotheses are set to be evaluated (in a later stage of this research project) in a group of persons with acute aphasia (within 3 months post-stroke), detect potential aberrant activation patterns (hence, inspect the integrity of the dorsal processing stream in acute stroke) and appraise possible clinical implementation of the proposed source reconstruction method.

Methods

Subjects

Forty-seven subjects (mean age 47.64 years $\pm$ 13.81), who were mainly recruited from hospital staff and senior club houses, participated in the study (24 female). All persons investigated were right-handed, as verified with the Dutch Handedness Inventory (DHI [39]). All the participants had Dutch as native language and reported to have normal hearing. None of them had neurological, psychiatric or speech and language developmental disorders, as verified by a questionnaire. At time of testing, none of the participants was on medication. The study was approved by the Ethics Committee of the University Hospital Ghent (Belgium) and an informed consent was obtained from all the subjects.

Paradigm and stimuli

An auditory phoneme discrimination (APD) task was used and consisted of three different auditory oddball paradigms both in a pre-attentive, passive and attentive, active condition. During the passive APD tasks, the subjects were instructed to ignore the stimuli and to focus their attention to a silent movie. To ensure that participants did not specifically pay attention towards the stimuli during the passive paradigm, they were asked questions about the movie afterwards, which they were informed about before the experiment started. During the active APD tasks, the subjects had to push a button every time they heard the infrequent stimulus. The sequence of the individual tasks was counterbalanced across subjects, meaning that every subject started with another task. Within the passive and active APD tasks, each block consisted of a total of 250 stimuli and 150 stimuli, respectively. The standard phoneme was [bə] and the deviant phonemes were [gə] (covering a PoA difference), [pə] (covering a voicing difference) and [mə] (covering a MoA difference) (Fig. 1). A summary of the acoustic–phonetic composition of every phoneme (analyzed with PRAAT [40]) is provided in Table 1. The APD tasks were created in such a way that the standard and deviant stimulus only differed in one phonemic contrast. Stimuli were generated using the Dutch website NeXTeNS (http://nextens.uvt.nl/demo.html) where text was converted to speech [41]. In all stimulus blocks, the standard and deviant phoneme appeared with a probability of 0.80 and 0.20, respectively. The stimuli were presented in a random order in which two deviants could not follow each other without having a standard in between. All the spoken phonemes had a duration of 150 ms. The interstimulus interval (ISI) was set at 500 ms
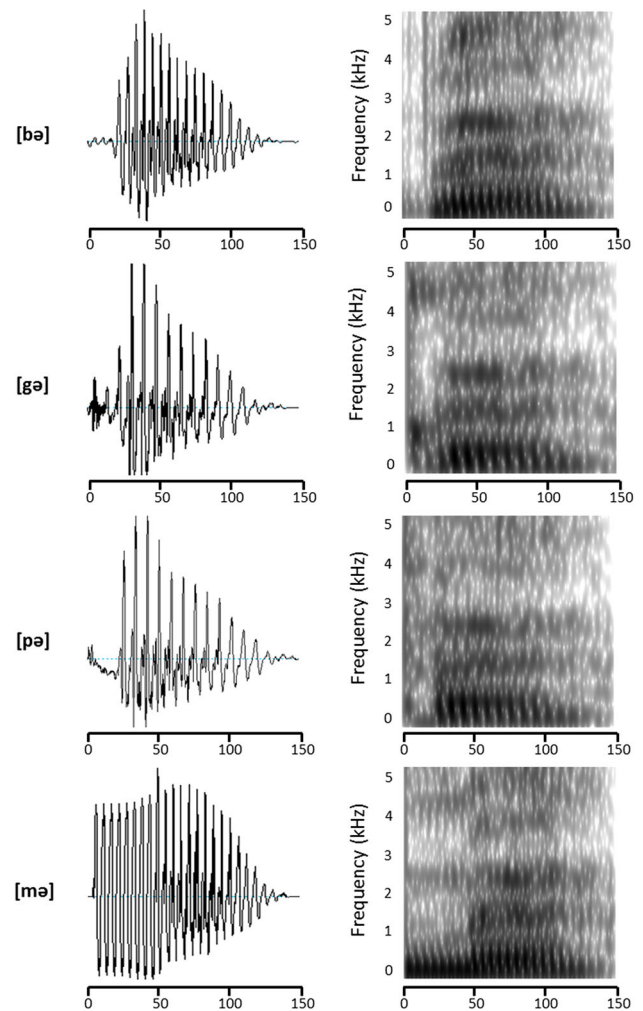

Fig. 1 Phoneme stimuli and their spectrograms. The *left*-side column represents the speech waveforms of every phoneme in the time domain (150 ms duration). The *right*-side column represents their spectrograms in the frequency domain, with time on the *x* axis and frequency (kHz) on the *y* axis

in the passive APD tasks and 2000 ms in the active APD tasks.

Electroencephalogram (EEG) recording and analysis

The EEG (0.5–100 Hz band-pass, notch filter disabled) was recorded through 23 Ag/AgCl-electrodes using a linked earlobes reference and an electrode placed on the forehead as ground. Electrodes were placed on the scalp according to the international 10–20 system. The impedance of the electrodes was kept below 5 kΩ. Data was collected using a SynAmp (Neuroscan) amplifier and was continuously digitized at a 500 Hz sampling rate.

EEG analysis was performed using BrainVision Analyzer 2 (Brain Products, Munich, Germany). According to

Table 1 Detailed acoustic–phonetic composition of the phoneme stimuli

	Standard [bə]	PoA deviant [gə]	Voicing deviant [pə]	MoA deviant [mə]
Voice-onset time (ms)	8	14	30	6
Formant frequencies (Hz)				
F0	151	135	136	152
F1	420	472	464	439
F2	1507	1435	1415	1476
F3	2331	2330	2318	2325
Formant transitions (Hz)				
F1				
Minimum	256	368	430	194
Maximum	417	450	469	421
%Increase	38.2	17.4	8.5	51.8
F2				
Minimum	1169	1135	1150	710
Maximum	1442	1375	1395	1406
%Increase	18.3	16.8	17.1	47.1
F3				
Minimum	2294	2152	2277	2318
Maximum	2336	2314	2352	2369
%Increase	1.8	6.9	3.2	2.2
Intensity (dB)				
Global	72.3	70.2	70.8	72.8
F1	48	46	45	46
dB re$_{mean phoneme level}$	−24.3	−24.2	−25.8	−26.8
F2	25	23	26	23
dB re$_{mean phoneme level}$	−47.3	−47.2	−44.8	−49.8
F3	24	22	22	19
dB re$_{mean phoneme level}$	−48.3	−48.2	−48.8	−53.8
Phonetic properties				
PoA	Bilabial	Velar	Bilabial	Bilabial
Voicing	Voiced	Voiced	Voiceless	Voiced
MoA	Oral	Oral	Oral	Nasal

Formant transitions were measured as the frequency change between minimum and maximum formant frequency (of F1, F2 and F3) and were calculated with respect to the median formant frequency of the "midvowel" (=percentage increase). It must be noted that formant transitions with the vowel [ə] are generally not large and, if present, are smaller with bilabial consonants than with velar consonants. The global intensity of all phonemes was as good as equal, with only a maximum difference of 2.6 dB. The intensity of F1, F2 and F3 was measured with reference to this global intensity value (dB re$_{mean phoneme level}$). The negative sign indicates that formant intensity in a limited frequency band (±50 Hz around median frequency) will always be smaller than the global intensity of the whole spectrum. Analyses were performed using PRAAT with the Burg-algorithm [40]

Ms milliseconds, *Hz* hertz, *F1* first formant, *F2* second formant, *F3* third formant, *dB* decibel, *re* reference, *PoA* place of articulation, *MoA* manner of articulation

our previous studies (e.g., reference [42]), the following procedure for preprocessing of EEG data was used: the EEG signal was filtered with a 0.5–30 Hz band-pass filter. Because no electrooculogram was applied, independent component analysis (ICA) had to be used to remove artifacts caused by eye movements by excluding two components based on inspection of the components' spatial distribution (prefrontal locations). These components comprised blinking and horizontal and vertical eye movements (saccades). For the three active APD tasks the EEG was segmented into 1100 ms long epochs from 100 ms pre-stimulus to 1000 ms post-stimulus. For the three passive APD tasks the EEG was segmented into 500 ms long epochs from 100 ms pre-stimulus to 400 ms post-stimulus.

The epochs were baseline corrected using a pre-stimulus window of 100 ms. All epochs containing artifacts exceeding 100 μV were rejected from further analysis, as were false responses in the active APD tasks. The data was converted to an average reference for further source reconstruction analysis. Standard and deviant trials were averaged separately. Difference waves were calculated only for the passive APD tasks to elicit the MMN by subtracting the average ERP of standard trials from the average ERP of deviant trials. Peak detection was carried out semi-automatically to measure peak latencies of the MMN in the passive APD tasks and peak latencies of the N100 and P300 in the active APD tasks. For the active APD tasks, reaction times and percentage of correct responses were calculated as well.

Source reconstruction and statistical analyses

All analyses were performed in the Statistical Parametric Mapping 8 software package (SPM 8: Welcome Department of Cognitive Neurology, University College, London, United Kingdom) implemented in MATLAB (the Math-Works, Inc., Massachusetts, USA).

First, a sensor-space analysis was executed to search for time points with significant differences in activation between phonemic contrasts PoA, voicing and MoA and the passive and active APD tasks. In the sensor-space analysis, the ERP data for each trial and each condition were converted into 3D images. These were generated by constructing 2D, 64 × 64 voxels resolution, scalp maps for each time point (using interpolation to estimate the activation between the electrodes) and by stacking the scalp maps over peristimulus time, resulting in [64 × 64 × number of time points]-images [43]. Using these images, statistical analyses were performed for each condition. Paired t tests were used to calculate two-tailed, two-sided (both activations AND deactivations) F-contrasts for the main effect of phonemic contrast (PoA vs voicing; voicing vs MoA; PoA vs MoA). The resulting statistical parametric map (SPM) was family-wise error (FWE) corrected for multiple comparisons using Random Field Theory [43–45], which takes into account the spatial correlation across voxels (i.e., that the tests are not independent). The sensor analysis was performed to search for maximal differences between conditions, so time windows were not chosen based on the peaks of the evoked waveforms for statistical analysis at the source level [46].

In a next step, source reconstruction was executed at the significant time points from the sensor analysis. The default 3-layered scalp-skull-brain template head model (the MNI brain) was used based on the Colin27 template [47], implemented in the SPM software. The default electrode positions were transformed to match the template head model, in which 8196 dipoles are assumed on a template cortical surface mesh. During the reconstruction, the coordinate system in which the electrode positions were originally represented was coregistered with the coordinate system of the template head model (i.e., MNI coordinates), again using the default electrode positions. In a next step, the forward model was calculated for each dipole on the cortical mesh, based on assumptions about the physical properties of the head, using the "bemcp" method (BEM) implemented in FieldTrip [48]. Finally, the actual inverse reconstruction was performed at group level for every condition over the whole ERP time window and was based on an empirical approach using the multiple sparse priors (MSP) algorithm [49]. Based on these reconstructions, 3D images were generated containing the evoked energy of the reconstructed activity in time windows corresponding with the significant time points from the sensor analysis. Using these images, additional second level analyses were performed to identify the most significant areas over subjects. Paired t tests were used to calculate one-tailed, one sided (activations OR deactivations) T-contrasts for the main effect of phonemic contrast (PoA vs voicing; voicing vs MoA; PoA vs MoA). Having corrected for multiple comparisons at the sensor-level, an uncorrected p value was set at 0.05 at the source level [46]. The resulting MNI coordinates holding significant activation differences between conditions, were explored by means of the SPM Anatomy toolbox developed by Eickhoff et al. [50] and were grouped to the following regions of interest: inferior frontal cortex (IFC), sensorimotor cortex (SMC), inferior parietal cortex (IPC) and superior temporal cortex (STC). It has been established in previous research that low-density recordings provide an accurate estimate of the underlying ERP generators, are sufficient to fully describe the variance of a typical ERP data set collected during an auditory oddball paradigm, when compared to high-density recordings, [51] and has been applied in neurolinguistic research before [52].

Additional statistical analyses were performed using IBM SPSS Statistics 22 on peak latencies in the passive and active APD tasks (MMN and N100/P300, respectively). For this, six clusters with the average latency of two electrodes were created, keeping midline electrodes separate: Anterior Left (F3, F7), Anterior Midline (Fz), Anterior Right (F4, F8), Central Left (T3, C3), Central Midline (Cz), Central Right (T4, C4), Posterior Left, (T5, P3) Posterior Midline (Pz), Posterior Right (T6, P4). Repeated measures ANOVA was carried out with factor contrasts (PoA vs. voicing vs. MoA), region (anterior vs. central vs. posterior) and laterality (left vs. midline vs. right). Reaction time was statistically analyzed as well, using repeated measures ANOVA with the factor contrasts (PoA vs. voicing vs. MoA). Greenhouse-Geisser correction

(GG) was applied when the assumption of sphericity was violated. Significance level was set at ≤0.05. Post hoc pairwise comparisons were computed using Bonferroni correction.

Results

ERP waveforms and behavioral results

A MMN was elicited in the passive APD task for all three phonemic contrasts, but with differences in morphology (Fig. 2). The MMN was significantly different from zero for all three phonemic contrasts, on left and right anterior, central and posterior locations $[T(46) = -9.64/-17.96, p < 0.001]$. In the active APD task an N100 and P300 was elicited with all three phonemic contrasts, but without clear differences in morphology (Fig. 3). With PoA as phonemic contrast, 97.65% of the deviant stimuli were identified, with voicing 96.8% and with MoA 98.34%. Subjects showed a mean response reaction time of 565.72 ms

(± 80.32) with PoA, 561.35 ms (± 94.52) with voicing and 535.44 ms (± 85.63) with MoA. There was a significant main effect for the phonemic contrasts $[F(2,90) = 5.78, p = 0.004]$. Faster response times occurred for the phonemic contrast MoA compared to PoA $(p = 0.001)$ and voicing $(p = 0.047)$.

Source reconstruction and ERP analyses

Passive APD task

Between 50 and 100 ms after stimulus presentation, the sensor-space analysis found significant differences between PoA and voicing $[F(1,46) = 28.43, p < 0.001, \text{FWE-corrected}]$, PoA and MoA $[F(1,46) = 85.76, p < 0.001, \text{FWE-corrected}]$ and MoA and voicing $[F(1,46) = 54.43, p < 0.001, \text{FWE-corrected}]$.

Second-level source analyses revealed the following pattern in the same time window: higher activation for MoA in sensorimotor regions compared to voicing $[T(46) = 4.28, p < 0.001]$ and PoA $[T(46) = 4.05, p < 0.001]$, higher
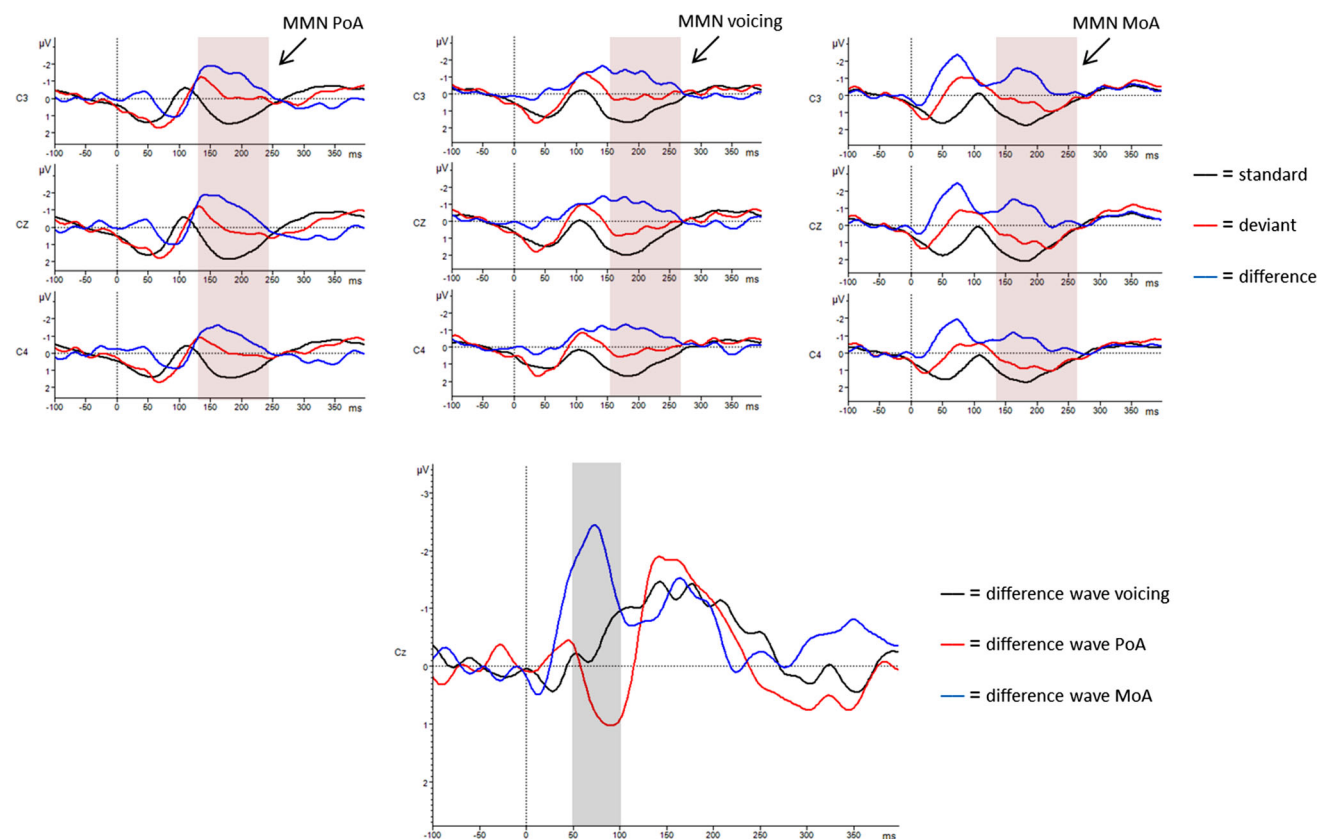

Fig. 2 Grand average ERPs for the passive APD task. In the *upper panel*, grand averaged ERPs are depicted for C3, Cz and C4 showing the response to the standard and deviant stimulus and the difference waveform (*black* standard; *red* deviant; *blue* difference waveform). The MMN is indicated in translucent *red*. In the *lower panel*, a detailed transcription is displayed at the Cz electrode with an overlay of the difference waves in response to the three phonemic contrasts (*black* voicing; *red* PoA; *blue* MoA). The time window (50–100 ms) which contains significant differences at source level during source reconstruction is indicated in translucent *blue*. Negative is plotted *upwards*

Fig. 3 Grand average ERPs for the active APD task. Grand averaged ERPs are depicted for P3, Pz and P4 with an overlay of the three phonemic contrasts (*black* voicing; *red* PoA; *blue* MoA) showing the response to the deviant stimulus. Three time windows (65–110 ms; 130–175 ms; 225–250 ms) containing significant differences at source level during source reconstruction are indicated in translucent *blue*. The last time window (310–340 ms) did not contain significantly different clusters at source level and is indicated in translucent *red*. Negative is plotted *upwards*

— = deviant voicing
— = deviant PoA
— = deviant MoA
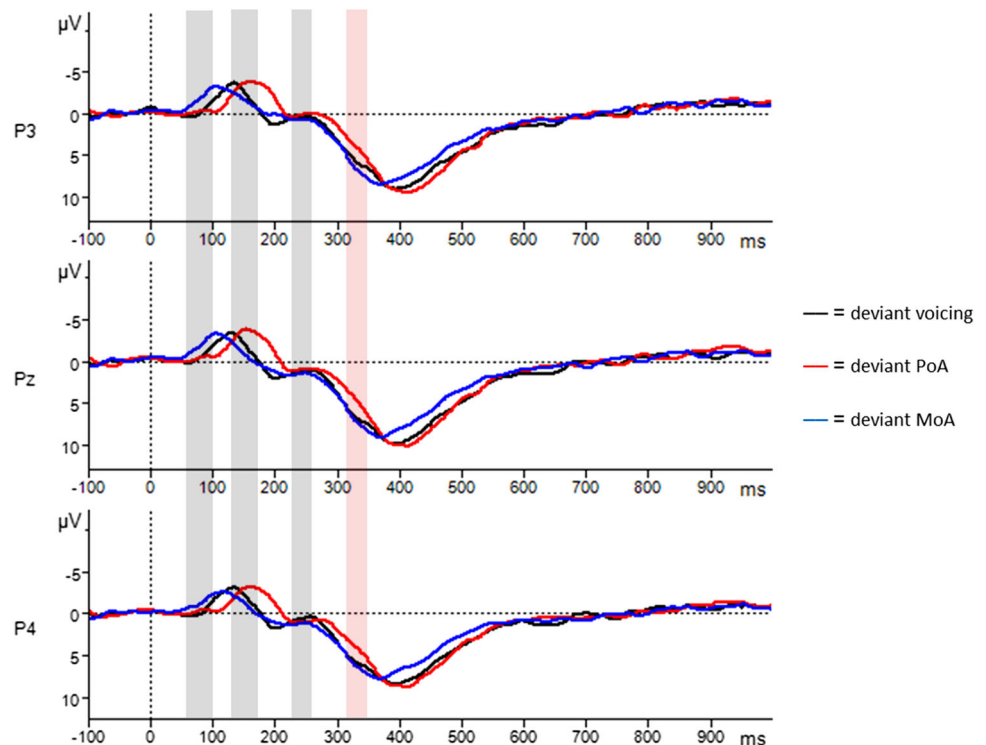
activation for MoA in inferior parietal regions compared to voicing [$T(46) = 2.79$, $p = 0.004$] and PoA [$T(46) = 3.05$, $p = 0.002$] and higher activation for MoA in superior temporal regions compared to voicing [$T(46) = 3.64$, $p < 0.001$] and PoA [$T(46) = 3.66$, $p < 0.001$]. No differences in activation were detected in any of the regions of interest when voicing and PoA were compared.

The time window with these activation differences (50–100 ms) corresponded to an epoch before the actual MMN onset and to a clear peak between 50 and 100 ms with MoA as phonemic contrast (Fig. 2). No significant timing differences between the phonemic contrasts were found at the level of MMN peak latency. Results of the source reconstruction can be found in Figs. 4 and 5 and Table 2.

Active APD task

The sensor-space analysis determined a first time window, between 65 and 110 ms, holding significant differences between PoA and voicing [$F(1,46) = 85.39$, $p < 0.001$, FWE-corrected], PoA and MoA [$F(1,46) = 112.10$, $p < 0.001$, FWE-corrected] and MoA and voicing [$F(1,46) = 41.02$, $p < 0.001$, FWE-corrected]. Source reconstruction showed that MoA elicited higher activation compared to PoA in inferior frontal [$T(46) = 2.25$, $p = 0.015$], sensorimotor [$T(46) = 4.85$, $p < 0.001$], inferior parietal [$T(46) = 3.94$, $p < 0.001$] and superior temporal regions [$T(46) = 4.27$, $p < 0.001$] and higher

activation compared to voicing in sensorimotor areas [$T(46) = 2.84$, $p = 0.003$]. In turn, voicing elicited higher activation in inferior frontal [$T(46) = 3.46$, $p = 0.001$], sensorimotor [$T(46) = 4.67$, $p < 0.001$], inferior parietal [$T(46) = 3.57$, $p < 0.001$] and superior temporal regions [$T(46) = 3.67$, $p < 0.001$] compared to PoA.

Further on, in a second time window, between 130 and 175 ms, the sensor-space analysis showed significant differences only between PoA and voicing [$F(1,46) = 90.26$, $p < 0.001$, FWE-corrected] and PoA and MoA [$F(1,46) = 91.11$, $p < 0.001$, FWE-corrected]. The difference between MoA and voicing disappeared. Moreover, during source reconstruction a reversed pattern emerged, compared to the 65 and 110 ms time-window, showing higher activation for PoA in inferior frontal regions compared to both voicing [$T(46) = 3.19$, $p = 0.001$] and MoA [$T(46) = 4.23$, $p < 0.001$], sensorimotor regions compared to both voicing [$T(46) = 3.79$, $p < 0.001$] and MoA [$T(46) = 3.90$, $p < 0.001$] and inferior parietal regions compared to both voicing [$T(46) = 3.12$, $p = 0.002$] and MoA [$T(46) = 2.94$, $p = 0.003$] with an additional higher activation in superior temporal regions for PoA compared to MoA [$T(46) = 2.52$, $p = 0.008$].

The established differences in this and the previous time window occurred around the N100 potential (Fig. 3) and were related to latency differences between phonemic contrasts. A region × contrasts interaction occurred in the statistical analysis on the N100 peak latency [$F(4,180) = 15.27$, $p < 0.001$, GG $\varepsilon = 0.68$], showing a
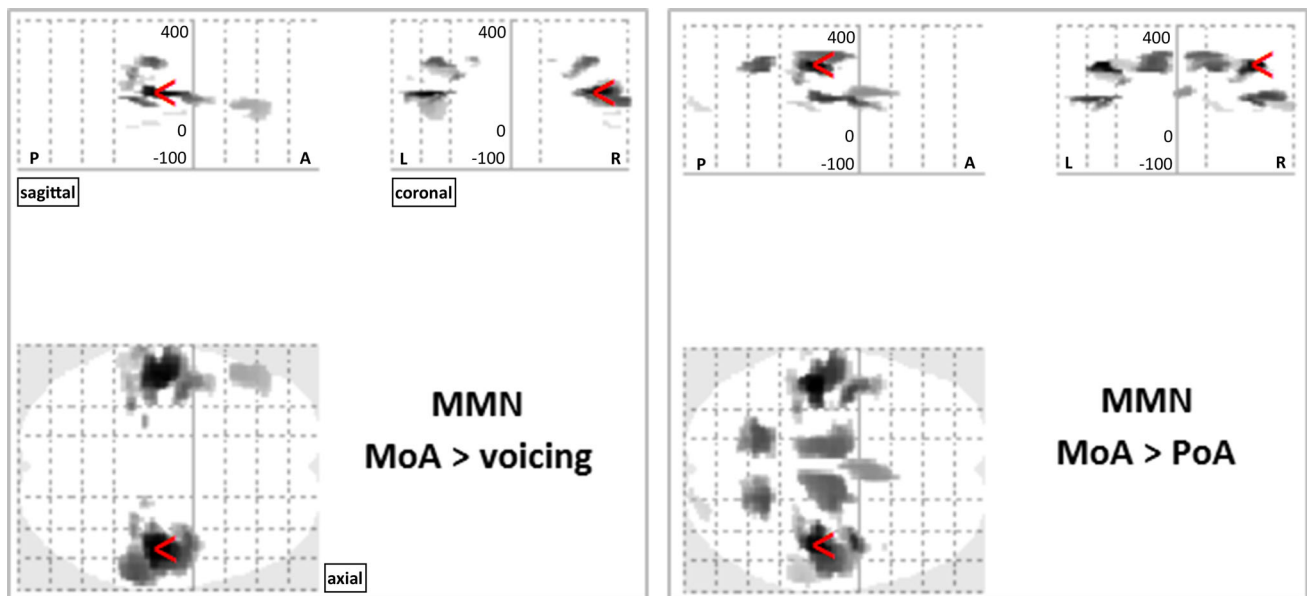
Fig. 4 Maximum intensity projections (MIP) of differences between phonemic contrasts during the passive APD task. The maximum intensity is projected on a glass brain in three orthogonal planes; axial (*bottom plot*), sagittal (*top left plot*) and coronal (*top right plot*). In the *top plots* of each box the time course of the region(s) with statistical differences in maximal activity is represented, with time going from the *bottom* of the plot (−100 ms) to the *top* (400 ms). The *bottom plot* shows the MIP at the time of the maximal activation. The MIP can be seen as an image in which all relevant activity above the $\alpha = 0.05$ (uncorrected) threshold is superimposed in a certain direction (axial, sagittal, and coronal). The *red arrow* indicates the global maximum activity. *P* posterior, *A* anterior, *L* left, *R* right
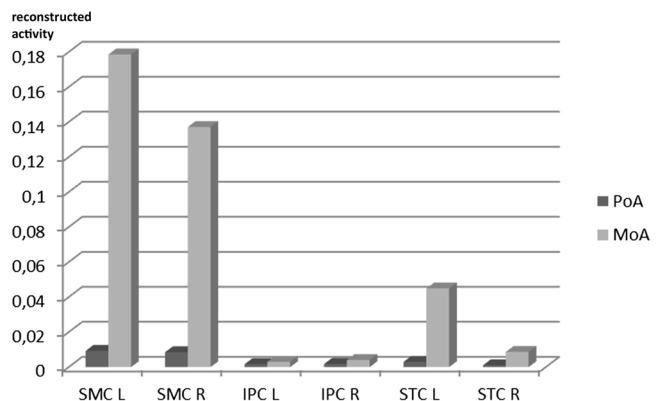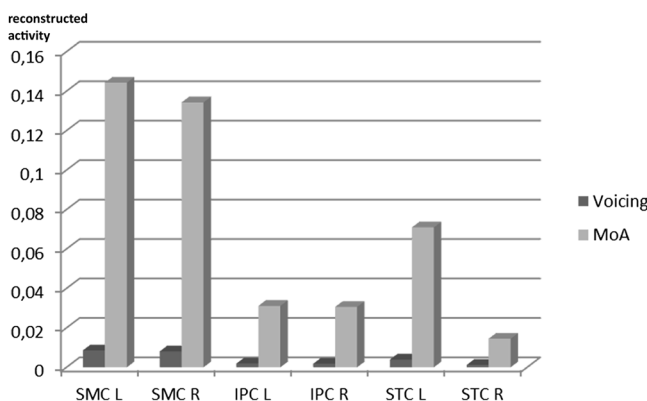

Fig. 5 Differences in activity levels between phonemic contrasts during the passive APD task. The y axis represents the amount of reconstructed activity per cluster (x axis), which is a number without unit. On the x axis the clusters of sensorimotor (SMC), inferior parietal (IPC) and superior temporal cortex (STC) are displayed. Between 50 and 100 ms pre-attentive phoneme discrimination based on MoA elicited more activation of sensorimotor, inferior parietal and superior temporal regions compared to voicing and PoA

longer latency for PoA compared to voicing ($p = 0.001$) and MoA ($p < 0.001$) and voicing compared to MoA ($p < 0.001$) in anterior regions, a longer latency for PoA ($p < 0.001$) and voicing ($p < 0.001$) compared to MoA in central regions and a longer latency for voicing compared to MoA ($p = 0.003$) in posterior regions. Because of these latency differences, possible differences in activation in the regions of interest between phonemic contrasts are searched for in different segments of the N100 potential (i.e. onset and peak).

Next, the sensor-space analysis defined a third (225–250 ms) time window which only contained significant differences between PoA and MoA [$F(1,46) = 36.18$, $p < 0.001$, FWE-corrected]. Higher activation of sensorimotor [$T(46) = 3.34$, $p = 0.001$] and superior temporal regions [$T(46) = 2.84$, $p = 0.003$] was elicited for PoA compared to MoA. A last time window (310–340 ms), corresponding to the onset of the P300 wave (Fig. 3), revealed significant differences between PoA and MoA during sensor-space analysis [$F(1,46) = 35.23$, $p < 0.001$, FWE-

Table 2 Results of the source reconstruction for the comparison between the phonemic contrasts PoA, voicing and MoA during passive, pre-attentive phoneme discrimination (MMN)

Time window	p	Cluster	Extent of voxels	Direction of difference
50–100 ms	<0.001	L sensorimotor cortex	1116	MoA > PoA
	<0.001	R sensorimotor cortex	1131	
	0.004	L inferior parietal cortex	80	
	0.006	R inferior parietal cortex	51	
	0.001	L superior temporal cortex	425	
	0.020	R superior temporal cortex	79	
	<0.001	L sensorimotor cortex	695	MoA > voicing
	<0.001	R sensorimotor cortex	725	
	0.002	L inferior parietal cortex	124	
	0.003	R inferior parietal cortex	612	
	0.004	L superior temporal cortex	36	
	0.001	R superior temporal cortex	288	

Reported are clusters of interest holding MNI coordinates with significant differences between phonemic contrasts, of which the extent of voxels is stated. Only the lowest p value within a significant cluster is displayed. The last column represents which phonemic contrast evoked the most energy

Ms millisecond, *p* level of significance, *L* left, *R* right, *PoA* place of articulation, *MoA* manner of articulation

corrected], though statistical analysis on source level could not confirm this. However, from the statistical analysis on the P300 peak latency a significant main effect for contrasts emerged, showing a longer latency for PoA compared to MoA [$F(2,90) = 12.15$, $p < 0.001$]. Results of the source reconstruction can be found in Figs. 6 and 7 and Table 3.

Discussion

We aimed to determine whether phoneme discrimination based on PoA elicits more motor cortical activity and whether phoneme discrimination based on voicing or MoA evokes more auditory cortical activity. Moreover, we investigated whether potential different cortical activation patterns related to these phonemic contrasts are organized in different temporal integration windows and whether increased attentional load would primarily be related to increased frontal activity. The current data show several time windows with significant differences in activation of auditory and motor regions associated with phoneme discrimination based on the three different phonemic contrasts and attentional load.

Already between 50 and 100 ms, pre-attentive phoneme discrimination with MoA as phonemic contrast elicits more activation in sensorimotor, inferior parietal and auditory superior temporal regions. Further on in time no dissimilarities arose between the different phonemic contrasts. Within the MoA-related difference waveform this time window corresponds to a pronounced peak before the definite MMN onset (Fig. 2). It seems that this pronounced pre-MMN peak represents the need for higher activation of auditory, temporal regions in conjunction with

sensorimotor regions during phoneme discrimination based on MoA, which seems to correspond to a dorsal processing stream typically involved in sublexical speech processing [53, 54]. Moreover, considering that this effect occurs before the actual MMN, it is possible that a preparatory phase before the actual phoneme discrimination process is needed, specifically when discrimination is dependent on MoA. So probably, during this preparatory phase a higher reliance on integration of both auditory features and motor based representations of the articulatory movements related to the phonemic contrast MoA is needed [7]. MoA is a phonemic contrast that has been rarely investigated, whereas PoA has been investigated extensively with respect to its representation in both auditory and motor regions [17, 27, 28, 55, 56]. Most of the studies implemented PoA as phonemic contrast most likely because of its clear-cut somatotopic correlates of the articulatory movements in the motor cortex (i.e. the lip/tongue area) [18, 55, 57]. This is less straightforward for MoA, as it determines phonemes from two perspectives: (1) nasality (height of the velum), which differentiates between oral ([bə]) and nasal ([mə]) phonemes; (2) duration or degree of constriction, which differentiates between plosives ([bə]) and fricatives ([və]). This possibly makes it more difficult to define proper subregions to investigate in the motor cortex. Nevertheless, the present study demonstrates that MoA compared to PoA and voicing reflects more auditory processing, probably related to the large formant transition of F1 and F2 (Table 1) due to deep resonances inherent to nasal sounds, and auditory-to-motor mapping in the run-up to phoneme discrimination, which partially corresponds to our first hypothesis (questioning whether phoneme discrimination based on PoA elicits more motor cortical
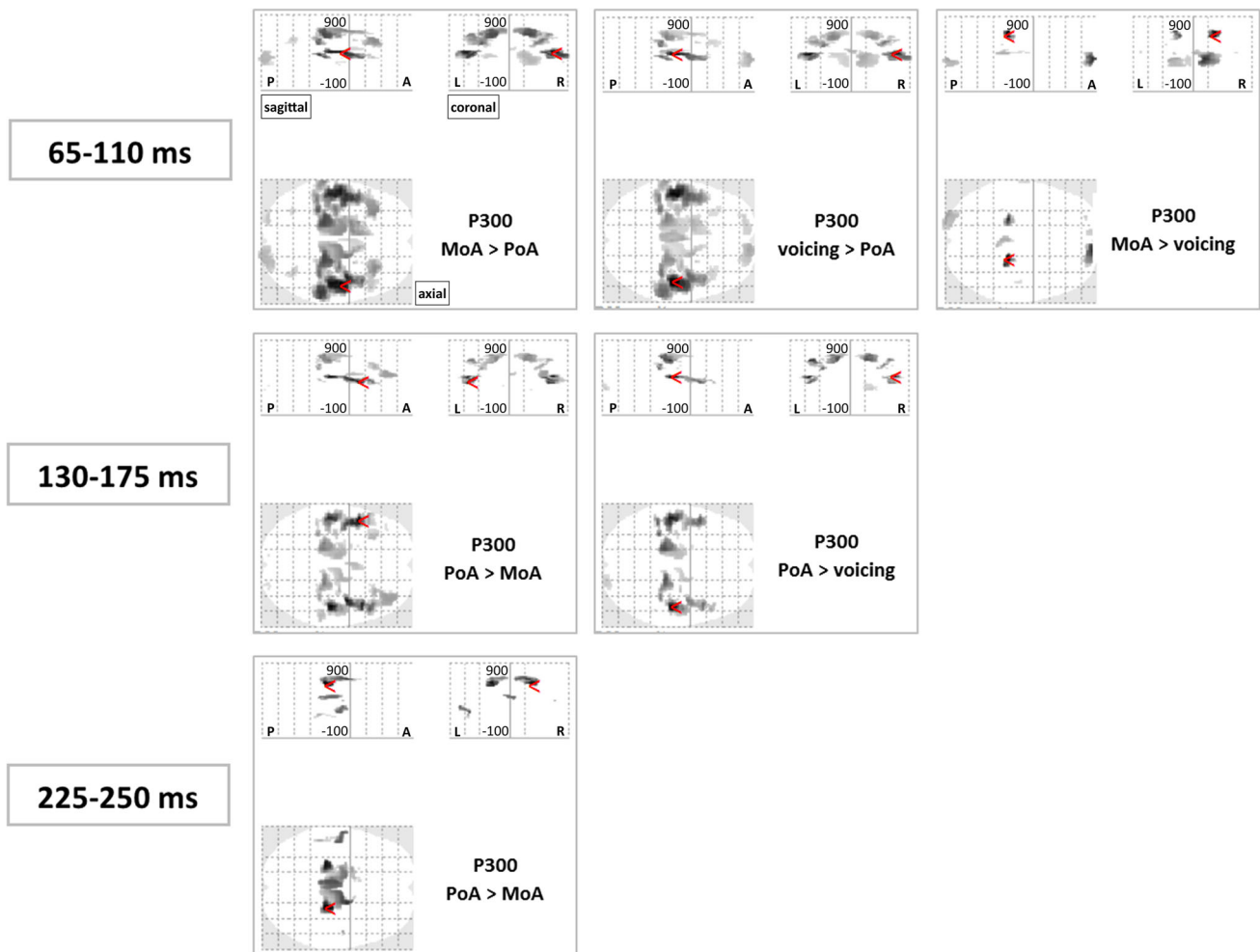
Fig. 6 Maximum intensity projections (MIP) of differences between phonemic contrasts during the active APD task. The maximum intensity is projected on a glass brain in three orthogonal planes; axial (*bottom plot*), sagittal (*top left plot*) and coronal (*top right plot*). In the *top plots* of each box the time course of the region(s) with statistical differences in maximal activity is represented, with time going from the *bottom* of the plot (−100 ms) to the *top* (900 ms). The *bottom plot* shows the MIP at the time of the maximal activation. The MIP can be seen as an image in which all relevant activity above the α = 0.05 (uncorrected) threshold is superimposed in a certain direction (axial, sagittal, and coronal). The *red arrow* indicates the global maximum activity. *P* posterior, *A* anterior, *L* left, *R* right

activity and whether phoneme discrimination based on voicing or MoA evokes more auditory cortical activity). Another interesting finding, not completely fulfilling our first hypothesis, is the fact that voicing did not show any difference with PoA and relied less on auditory cortical activity than MoA. Compared to MoA, voicing has more delineated motor features (related to the laryngeal motor function), which already proved to be supportive for discrimination of vocal pitch [55]. Perhaps this explains the lack of difference between voicing and PoA in motor regions (equal reliance on motor features) and auditory regions (lack of increased auditory processing for voicing) and the existing difference between MoA and voicing in auditory regions (less reliance on auditory features for voicing, when compared to MoA, than hypothesized).

During attentive phoneme discrimination, a spatial differentiation starts already between 65 and 110 ms, just as during the pre-attentive condition, where motor areas, inferior parietal regions and auditory superior temporal regions are more activated during phoneme discrimination based on MoA and voicing. Later on, between 130 and 175 ms a reversed pattern emerged showing more activation of exactly the same motor and auditory regions during phoneme discrimination based on PoA, instead of MoA and voicing. Finally, between 225 and 250 ms only the difference between PoA and MoA remained, showing more sensorimotor and superior temporal activation during phoneme discrimination based on PoA. Based on the underlying ERP waveforms, it appears that in the early time window (65–110 ms) an initiation of an N100 is not
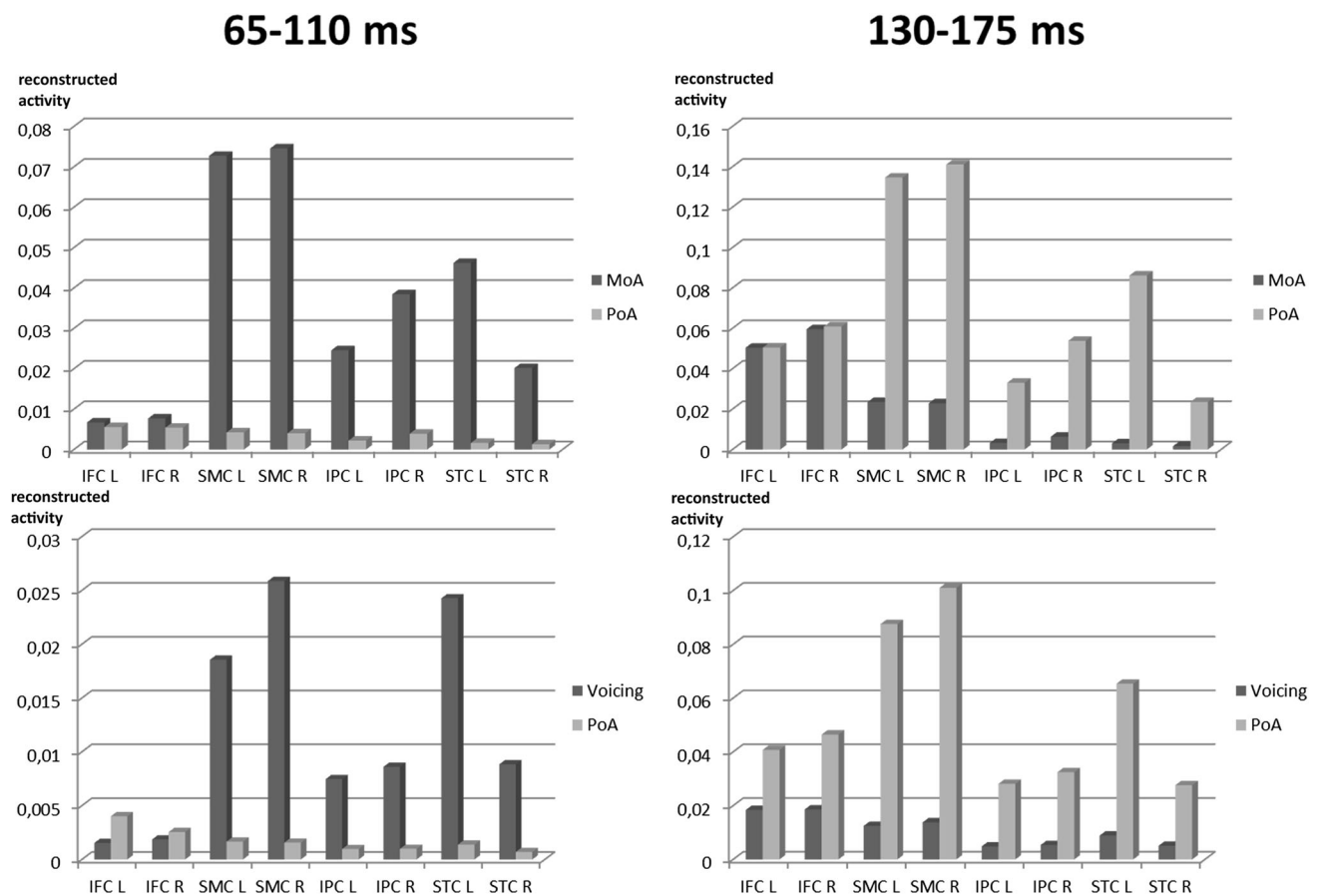
65-110 ms

reconstructed activity

130-175 ms

reconstructed activity

reconstructed activity

reconstructed activity

Fig. 7 Differences in activity levels between phonemic contrasts during the active APD task represented in the first two time windows. The *y* axis represents the amount of reconstructed activity per cluster (*x* axis), which is a number without unit. On the *x* axis the inferior frontal (IFC), sensorimotor (SMC), inferior parietal (IPC) and superior temporal cortex (STC) are shown. Between 65 and 110 ms

attentive phoneme discrimination based on MoA and voicing requires more activation of inferior frontal, sensorimotor, inferior parietal and superior temporal regions. Between 130 and 175 ms a reversed pattern emerged showing more involvement of the same motor and auditory areas during phoneme discrimination based on PoA

yet present with PoA as phonemic contrast while this is the case for MoA and voicing. Moreover, the following time frame (130–175 ms) corresponds to the N100 onset of PoA while the N100 in response to MoA and voicing is already in its offset phase. These timing differences around the N100 potential are supported by statistically significant differences between phonemic contrasts in N100 peak latency (showing a longer latency for PoA compared to voicing and MoA). The N100 potential is generally related to primary analyses, preparatory processes like detection and identification of phonemes based on their different constituent features [27, 31, 58]. So, the preparation for phoneme discrimination based on PoA occurs later than the preparation for phoneme discrimination based on MoA and voicing, which clearly illustrates the importance of taking temporal processing into account. Nonetheless, the same structures are engaged, showing auditory-motor integration during these preparatory processes, seemingly rejecting the hypothesis of more auditory cortical activity for MoA and/

or voicing. However, in the comparison between PoA and voicing the latter relies on both auditory temporal and motor regions between 65 and 110 ms, whereas the former does not and depends only on motor areas in the later time window between 130 and 175 ms. Hence, PoA does appear to be more motor imprinted, most likely due to its distinctive articulatory gestures, while voicing is more acoustically imprinted in the brain, probably related to its distinct rapid temporal cues (voice-onset time), at least during active, attentive phoneme discrimination (as there is more motor involvement during passive phoneme discrimination, see above).

A last finding is the absence of regional activation differences between any of the phonemic contrasts during pre-attentive or attentive phoneme discrimination in the time window associated with the MMN peak or onset of the P300 peak (between 310 and 340 ms). However, in the attentive condition there were significant longer latencies for PoA compared to MoA, showing a similar pattern as the

Table 3 Results of the source reconstruction for the comparison between the phonemic contrasts PoA, voicing and MoA during active, attentive phoneme discrimination (P300)

Time window	p	Cluster	Extent of voxels	Direction of difference
65–110 ms	0.029	L inferior frontal cortex	24	MoA > PoA
	<0.001	L sensorimotor cortex	1697	
	<0.001	R sensorimotor cortex	1736	
	<0.001	L inferior parietal cortex	157	
	0.001	R inferior parietal cortex	173	
	0.001	L superior temporal cortex	78	
	<0.001	R superior temporal cortex	197	
	0.001	L inferior frontal cortex	529	Voicing > PoA
	<0.001	L sensorimotor cortex	1366	
	<0.001	R sensorimotor cortex	1242	
	0.001	L inferior parietal cortex	149	
	0.001	R inferior parietal cortex	119	
	0.003	L superior temporal cortex	61	
	0.001	R superior temporal cortex	205	
	0.011	L sensorimotor cortex	93	MoA > voicing
	0.007	R sensorimotor cortex	129	
130–175 ms	<0.001	L inferior frontal cortex	398	PoA > MoA
	<0.001	L sensorimotor cortex	773	
	<0.001	R sensorimotor cortex	876	
	0.005	L inferior parietal cortex	30	
	0.008	R inferior parietal cortex	31	
	0.015	R superior temporal cortex	32	
	0.003	L inferior frontal cortex	243	PoA > voicing
	0.001	L sensorimotor cortex	703	
	<0.001	R sensorimotor cortex	573	
	0.003	L inferior parietal cortex	49	
	0.006	R inferior parietal cortex	35	
225–250 ms	0.002	L sensorimotor cortex	362	PoA > MoA
	0.002	R sensorimotor cortex	546	
	0.007	L superior temporal cortex	65	

Reported are clusters of interest holding MNI coordinates with significant differences between phonemic contrasts, of which the extent of voxels is stated. Only the lowest p value within a significant cluster is displayed. The last column represents which phonemic contrast evoked the most energy

Ms millisecond, *p* level of significance, *L* left, *R* right, *PoA* place of articulation, *MoA* manner of articulation

timing differences around the N100 potential as mentioned above. So, attentive phoneme discrimination with PoA remains to be prolonged compared to MoA. Apparently, in contrast with the phonological preparatory processes, the actual phoneme discrimination process does not evoke an increased activation of auditory-motor regions depending on the phonemic contrast, despite the timing differences. In addition, studies which found differences between articulatory representations in the motor cortex mainly used passive listening to phonemes, merely probing phoneme perception and recognition of one single phonemic contrast [16–18, 20]. Perhaps purely discriminating between phonemes does not lead to altered activity in auditory and

motor regions with respect to different phonemic contrasts whereas the preparatory phases, such as phoneme detection and identification, do so. The fact that the auditory-motor integration during the preparatory phases for PoA phoneme discrimination occur later than the auditory-acoustic processes as preparation for MoA/voicing phoneme discrimination, can perhaps be explained, although still speculative, from a developmental perspective. Already in utero, the foetus is capable of hearing environmental sounds, considering that the auditory organ reaches its adult size by the twentieth week of gestation [59]. This is supported by the presence of auditory evoked responses in the foetal brain as measured with foetal magnetoencephalography (fMEG)

[60]. Such auditory evoked magnetic fields have proven that foetuses, mostly between 27 and 36 weeks gestational age, are capable of detecting tone-frequency changes [61–64] and that neonates can process rapid temporal cues within 3 weeks after birth [60]. Evidently, processing of sounds during foetal development and early infancy highly relies on acoustic-auditory features. As already mentioned in the introduction, the phonemic contrasts voicing and MoA are highly acoustically founded and well-defined (temporal cue of voice onset time and F2 transition related to deep resonances, respectively) whereas PoA has a more elusive acoustic nature (lack of specific, invariant acoustic cues and presence of multiple acoustic variables; [36]) with a marked audio-visual, motor-phonetic component. The motor features are learned and related to auditory signals well after birth, when the infant starts babbling and receives tactile, proprioceptive and auditory feedback when producing sounds and sees and hears sounds being produced [65]. Because of the vague acoustic composition of PoA, an integrative perception strategy must be engaged based on auditory and motor features, whereas perception of voicing and MoA can be completed successfully based on auditory features only. It is possible that the early auditory-acoustic imprinting of voicing and MoA and the later auditory-motor integration for PoA during phoneme perception found in the last study might be a reflection of a very early, prenatal development of the central auditory system and a later, postnatal formation of the auditory-motor integration network. The fact that this (spatio)temporal dissociation of the preparatory processes for PoA and MoA/voicing only surfaces during the attentive phoneme discrimination task, can potentially be related to the explicit conscious cognitive process underlying the nature of this active oddball paradigm, opening a 'gate' for uncovering long-established developmental patterns (which is less possible with automatic, pre-attentive tasks).

In any way, it is demonstrated that starting from 50 ms after phoneme presentation, activity of inferior frontal, sensorimotor, inferior parietal and auditory regions (corresponding to a dorsal processing stream) temporally differentiates between the three phonemic contrasts during phoneme discrimination.

An effect of attention was observed when the different phonemic contrasts were compared in the pre-attentive and attentive condition. Activity differences between phonemic contrasts were found in the inferior frontal region during attentive phoneme discrimination, whereas no such dissimilarities were found during pre-attentive phoneme discrimination. Frontal regions have repeatedly been associated with selective and directed attention processes [66–68]. Moreover, a study where the MMN was compared in a conventional, pre-attentive condition and a more attention-demanding design attributed greater inferior frontal involvement during the attention-demanding design to some form of "attention-switching towards auditory linguistic input" [34]. With respect to the present results in the attentive condition, this might indicate that a voluntary switch of attention to the heard phonemes occurs in an early time window for MoA and voicing and later in time, between 130 and 175 ms, for PoA. This might also explain why the PoA-related preparatory analyses start later in time than the MoA-related and voicing-related preparatory processes.

To conclude, pre-attentive as well as attentive auditory phoneme discrimination demonstrated a clear differentiation between all three phonemic contrasts. Firstly, during the early, preparatory phases (50–100 ms) more intense auditory processing and auditory-to-motor mapping was required for pre-attentive phoneme discrimination based on MoA. Secondly, there was a delayed initiation of the preparatory phases with the phonemic contrast PoA, though the same auditory and motor areas were engaged as the other two phonemic contrasts (65–110 ms; 130–175 ms), during attentive phoneme discrimination. A larger fronto-parietal attention-related network was active during attentive (voluntary) as well as pre-attentive (involuntary) phoneme discrimination, but showed greater variation in time and degree of influence for the phonemic contrast PoA. Above that, with the phonemic contrasts MoA and voicing there was a higher emphasis on auditory regions during involuntary attention-allocation, whereas for PoA attention-allocation relied solely on motor areas. Finally, it must be taken into account that the present study is limited in its spatial sampling because of the less dense electrode coverage. As such, conclusions regarding location of activation differences must be interpreted with due caution. However, we do believe this source reconstruction method is worth evaluating in a population of persons with acute aphasia to assess sublexical processing on a neuronal level.

References

1. Friederici AD (2011) The brain basis of language processing: from structure to function. Physiol Rev 91:1357–1392

2. Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393–402

3. Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry M-S, Umarova R, Musso M, Glauche V, Abel S (2008) Ventral and dorsal pathways for language. Proc Natl Acad Sci 105:18035–18040

4. Binder J, Frost J, Hammeke T, Bellgowan P, Springer J, Kaufman J, Possing E (2000) Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 10:512–528

5. DeWitt I, Rauschecker JP (2012) Phoneme and word recognition in the auditory ventral stream. Proc Natl Acad Sci 109:E505–E514

6. Friederici AD (2012) The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn Sci 16:262–268

7. Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92:67–99

8. Jäncke L, Wüstenberg T, Scheich H, Heinze H-J (2002) Phonetic perception and the temporal cortex. Neuroimage 15:733–746

9. Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. Cereb Cortex 11:946–953

10. Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:389–395

11. Burton MW (2009) Understanding the role of the prefrontal cortex in phonological processing. Clin Linguist Phon 23:180–195

12. Friederici AD (2009) Pathways to language: fiber tracts in the human brain. Trends Cogn Sci 13:175–181

13. Celsis P, Boulanouar K, Doyon B, Ranjeva J, Berry I, Nespoulous J, Chollet F (1999) Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. Neuroimage 9:135–144

14. Hickok G, Buchsbaum B, Humphries C, Muftuler T (2003) Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J Cogn Neurosci 15:673–682

15. Turkeltaub PE, Branch Coslett H (2010) Localization of sublexical speech perception components. Brain Lang 114:1–15

16. D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L (2009) The motor somatotopy of speech perception. Curr Biol 19:381–385

17. Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M (2007) The essential role of premotor cortex in speech perception. Curr Biol 17:1692–1696

18. Pulvermüller F, Huss M, Kherif F, del Prado Martin FM, Hauk O, Shtyrov Y (2006) Motor cortex maps articulatory features of speech sounds. Proc Natl Acad Sci 103:7865–7870

19. Pulvermüller F, Fadiga L (2010) Active perception: sensorimotor circuits as a cortical basis for language. Nat Rev Neurosci 11:351–360

20. Wilson SM, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. Nat Neurosci 7:701–702

21. Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T (1997) Human brain language areas identified by functional magnetic resonance imaging. J Neurosci 17:353–362

22. Kotz SA, Schwartze M (2010) Cortical speech processing unplugged: a timely subcortico-cortical framework. Trends Cogn Sci 14:392–399

23. Burton MW (2001) The role of inferior frontal cortex in phonological processing. Cogn Sci 25:695–709

24. Londei A, D'Ausilio A, Basso D, Sestieri C, Gratta CD, Romani GL, Belardinelli MO (2010) Sensory-motor brain network connectivity for speech comprehension. Hum Brain Mapp 31:567–580

25. Zaehle T, Geiser E, Alter K, Jancke L, Meyer M (2008) Segmental processing in the human auditory dorsal stream. Brain Res 1220:179–190

26. Diesch E, Luce T (1997) Magnetic fields elicited by tones and vowel formants reveal tonotopy and nonlinear summation of cortical activation. Psychophysiology 34:501–510

27. Obleser J, Lahiri A, Eulitz C (2003) Auditory-evoked magnetic field codes place of articulation in timing and topography around 100 milliseconds post syllable onset. Neuroimage 20:1839–1847

28. Obleser J, Lahiri A, Eulitz C (2004) Magnetic brain response mirrors extraction of phonological features from spoken vowels. J Cogn Neurosci 16:31–39

29. Obleser J, Boecker H, Drzezga A, Haslinger B, Hennenlotter A, Roettinger M, Eulitz C, Rauschecker JP (2006) Vowel sound extraction in anterior superior temporal cortex. Hum Brain Mapp 27:562–571

30. Tavabi K, Obleser J, Dobel C, Pantev C (2007) Auditory evoked fields differentially encode speech features: an MEG investigation of the P50 m and N100m time courses during syllable processing. Eur J Neurosci 25:3155–3162

31. Cooper RJ, Todd J, McGill K, Michie PT (2006) Auditory sensory memory and the aging brain: a mismatch negativity study. Neurobiol Aging 27:752–762

32. Linden DEJ (2005) The P300: where in the brain is it produced and what does it tell us? Neuroscientist 11:563–576

33. Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, Allik J, Sinkkonen J, Alho K (1997) Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 385:432–434

34. Shtyrov Y, Kujala T, Pulvermüller F (2009) Interactions between language and attention systems: early automatic lexical processing? J Cogn Neurosci 22:1465–1478

35. Möttönen R, Watkins KE (2009) Motor representations of articulators contribute to categorical perception of speech sounds. J Neurosci 29:9819–9825

36. Sinnott JM, Gilmore CS (2004) Perception of place-of-articulation information in natural speech by monkeys versus humans. Percept Psychophys 66:1341–1350

37. MacGregor LJ, Pulvermüller F, van Casteren M, Shtyrov Y (2012) Ultra-rapid access to words in the brain. Nat Commun 3:711

38. Pulvermüller F, Shtyrov Y, Ilmoniemi R (2003) Spatiotemporal dynamics of neural language processing: an MEG study using minimum-norm current estimates. Neuroimage 20:1020–1025

39. Van Strien JV (2003) Vragenlijst voor handvoorkeur. Nederlands Tijdschrift voor de Psychologie en Haar Grensgebieden 47:88–92

40. Boersma P, Weenink D (2010) Praat: doing phonetics by computer [Computer program]. Version 5.1.45. http://www.praat.org/. Retrieved 26.10.10

41. Marsi E, Busser B, Daelemans W, Hoste V, Reynaert M (2002) Combining information sources for memory-based pitch accent placement. Paper presented at the International Conference on Spoken Language Processing, Denver Co

42. Aerts A, van Mierlo P, Hartsuiker RJ, Hallez H, Santens P, De Letter M (2013) Neurophysiological investigation of phonological input: aging effects and development of normative data. Brain Lang 125(3):253–263

43. Litvak V, Mattout J, Kiebel S, Phillips C, Henson R, Kilner J, Barnes G, Oostenveld R, Daunizeau J, Flandin G (2011) EEG and MEG data analysis in SPM8. Comput Intell Neurosci 2011, Article ID 852961:32. doi:10.1155/2011/852961

44. Kiebel SJ, Friston KJ (2004) Statistical parametric mapping for event-related potentials (II): a hierarchical temporal model. Neuroimage 22:503–520

45. Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996) A unified statistical approach for determining

significant signals in images of cerebral activation. Hum Brain Mapp 4:58–73

46. Gross J, Baillet S, Barnes GR, Henson RN, Hillebrand A, Jensen O, Jerbi K, Litvak V, Maess B, Oostenveld R, Parkkonen L, Taylor JR, van Wassenhove V, Wibral M, Schoffelen JM (2013) Good practice for conducting and reporting MEG research. Neuroimage 65:349–363

47. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, Evans AC (1998) Design and construction of a realistic digital brain phantom. IEEE Trans Med Imaging 17:463–468

48. Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:156869

49. Friston K, Harrison L, Daunizeau J, Kiebel S, Phillips C, Trujillo-Barreto N, Henson R, Flandin G, Mattout J (2008) Multiple sparse priors for the M/EEG inverse problem. Neuroimage 39:1104–1120

50. Eickhoff SB, Paus T, Caspers S, Grosbras M-H, Evans AC, Zilles K, Amunts K (2007) Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36:511–521

51. Kayser J, Tenke CE (2006) Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. Clin Neurophysiol 117:369–380

52. Vanhoutte S, Strobbe G, van Mierlo P, Cosyns M, Batens K, Corthals P, De Letter M, Van Borsel J, Santens P (2015) Early lexico-semantic modulation of motor related areas during action and non-action verb processing. J Neurolinguist 34:65–82

53. Chevillet MA, Jiang X, Rauschecker JP, Riesenhuber M (2013) Automatic phoneme category selectivity in the dorsal auditory stream. J Neurosci 33:5208–5215

54. Dehaene-Lambertz G, Hertz-Pannier L, Dubois J (2006) Nature and nurture in language acquisition: anatomical and functional brain-imaging studies in infants. Trends Neurosci 29:367–373

55. D'Ausilio A, Bufalari I, Salmas P, Busan P, Fadiga L (2011) Vocal pitch discrimination in the motor system. Brain Lang 118:9–14

56. Möttönen R, Dutton R, Watkins KE (2013) Auditory-motor processing of speech sounds. Cereb Cortex 23:1190–1197

57. Bouchard KE, Mesgarani N, Johnson K, Chang EF (2013) Functional organization of human sensorimotor cortex for speech articulation. Nature 495:327–332

58. Näätänen R, Kujala T, Winkler I (2011) Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. Psychophysiology 48:4–22

59. Bibas AG, Xenellis J, Michaels L, Anagnostopoulou S, Ferekidis E, Wright A (2008) Temporal bone study of development of the organ of Corti: correlation between auditory function and anatomical structure. J Laryngol Otol 122(4):336–342

60. Sheridan CJ, Matuz T, Draganova R, Eswaran H, Preissl H (2010) Fetal magnetoencephalography—achievements and challenges in the study of prenatal and early postnatal brain responses: a review. Infant Child Dev 19:80–93

61. Draganova R, Eswaran H, Murphy P, Huotilainen M, Lowery C, Preissl H (2005) Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. NeuroImage 28:354–361

62. Draganova R, Eswaran H, Murphy P, Lowery C, Preissl H (2007) Serial magnetoencephalographic study of fetal and newborn auditory discriminative evoked responses. Early Hum Dev 83:199–207

63. Holst M, Eswaran H, Lowery C, Murphy P, Norton J, Preissl H (2005) Development of auditory evoked fields in human fetuses and newborns: a longitudinal MEG study. Clin Neurophysiol 116:1949–1955

64. Huotilainen M, Kujala A, Hotakainen M, Parkkonenm L, Taulu S, Simola J et al (2004) Short-term memory functions of the human fetus recorded with magnetoencephalography. NeuroReport 16(1):81–84

65. Guenther FH, Ghosh SS, Tourville JA (2006) Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang 96:280–301

66. Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci 23:475–483

67. Fuster JM (2000) Executive frontal functions. Exp Brain Res 133:66–70

68. Peers PV, Ludwig CJ, Rorden C, Cusack R, Bonfiglioli C, Bundesen C, Driver J, Antoun N, Duncan J (2005) Attentional functions of parietal and frontal cortex. Cerebral Cortex 15:1469–1484