

## Electrophysiological responses to emotional prosody perception in cochlear implant users<sup>☆</sup>



D. Agrawal<sup>a,\*</sup>, J.D. Thorne<sup>b</sup>, F.C. Viola<sup>b</sup>, L. Timm<sup>a</sup>, S. Debener<sup>b</sup>, A. Büchner<sup>c</sup>, R. Dengler<sup>a</sup>, M. Wittfoth<sup>a,d</sup>

<sup>a</sup> Department of Neurology, Hannover Medical School, Hannover, Germany

<sup>b</sup> Department of Psychology, Carl von Ossietzky Universität, Oldenburg, Germany

<sup>c</sup> Department of Otolaryngology, Hannover Medical School, Hannover, Germany

<sup>d</sup> NICA (Neuroimaging and Clinical Applications), Hannover, Germany

### ARTICLE INFO

#### Article history:

Received 22 August 2012

Received in revised form 31 December 2012

Accepted 3 January 2013

Available online 14 January 2013

#### Keywords:

Emotional prosody

Cochlear implants

ERP

P200

Gamma band power

### ABSTRACT

The present electroencephalographic (EEG) study investigated the ability of cochlear implant (CI) users to recognize emotional prosody. Two CI speech-processing strategies were compared: the ACE (Advance Combination Encoder) and the newly developed MP3000. Semantically neutral sentences spoken in three different emotional prosodies (neutral, angry, happy) were presented to 20 post-lingually deafened CI users and age-matched normal-hearing controls. Event related potentials (ERPs) were recorded to study the N100 and the P200 responses. In addition, event-related spectral power modulations were calculated to study the brain activity corresponding to the recognition of prosody in earlier (0–400) as well as later (600–1200) part of the stimuli where the prosodic features differed maximally. CI users with MP3000 strategy showed a higher proportion of correctly recognized prosodic information compared to the ACE strategy users. Our ERP results demonstrated that emotional prosody elicited significant N100 and P200 peaks. Furthermore, the P200 amplitude in response to happy prosodic information was significantly more positive for the MP3000 strategy compared to the ACE strategy. On spectral power analysis, two typical gamma activities were observed in the MP3000 users only: (1) an early gamma activity in the 100–250 ms time window reflecting bottom-up attention regulation; and (2) a late gamma activity between 900 and 1100 ms post-stimulus onset, probably reflecting top-down cognitive control. Our study suggests that the MP3000 strategy is better than ACE in regard to happy prosody perception. Furthermore, we show that EEG is a useful tool that, in combination with behavioral analysis, can reveal differences between two CI processing strategies for coding of prosody-specific features of language.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

### 1. Introduction

In spoken language, emotionally salient information can be provided by variations in speech melody (emotional prosody) or by emotional semantics (verbal emotional content). Emotional prosody is the ability to express emotions through variations of pitch, intensity and duration (Scherer, 2003). Individuals with severe to profound hearing loss have a limited dynamic range in these parameters, thus affecting prosody recognition (Hopyan et al., 2011). Cochlear implants (CI) are thought to improve not only language perception per se, but also specific aspects of language. However, it is still an open question whether such improvements include the recognition of emotional prosodic information.

The CI encodes sounds electronically, bypassing the damaged cochlea, and electrically stimulates the auditory nerve. Speech coding strategies are extremely important in CI processing as they decompose audio signals into different frequency bands and deliver the stimulation pattern to electrodes, maximizing the user's communicative potential. A number of speech processing strategies have been developed over the past two decades mimicking firing patterns inside the normal cochlea (Loizou, 1999). The Advance Combination Encoder (ACE) was developed in the 1990s. This strategy separates speech signals into a number of sub bands (M) and derives the envelope information from each band signal. A subset of these (N) with the largest amplitude is then selected for stimulation (N out of M). In 2005, a new strategy, Psychoacoustic Advanced Combination Encoder (PACE), commercially known as MP3000 (this term will be used here), was developed. This strategy is based on a psychoacoustic-masking model that neglects redundant signals, thus saving valuable bandwidth for those components that are usually perceived by normal-hearing (NH) individuals. The strategy is similar to the MP3 compression algorithm (Nogueira et al., 2005).

There are reports highlighting the performance of CI users in understanding speech, in particular from phoneme to sentences, in quiet as well as in noisy environments. For example, researchers (Fu

<sup>☆</sup> This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

\* Corresponding author at: Department of Neurology, Hannover Medical School, Carl-Neuberg Strasse 1, Hannover, Germany. Tel.: +49 511 532 3737; fax: +49 511 532 3115.

E-mail address: [Agrawal.deepashri@mh-hannover.de](mailto:Agrawal.deepashri@mh-hannover.de) (D. Agrawal).

et al., 2005) reported that CI users' voice gender identification was nearly perfect (94% correct) when large differences in the fundamental frequency (F0) existed between male and female talkers. House and colleagues used semantically neutral Swedish utterances of four target emotions (angry, happy, sad, and neutral), and found that for CI users, the performance of mean vocal emotion recognition increased from 44% to 51% one year after processor activation (House, 1994). Comparably, another study (Luo and Fu, 2007) investigated the ability of NH listeners and CI users to recognize vocal emotions. CI users performed poorly with their own conventional processor, but their performance significantly improved as the number of channels was increased. Taken together these studies reflect mixed results, which might be due to the lack of objective dependent variables used. To overcome this issue event-related potentials (ERPs) could be used, as they do not rely on subjective and behavioral output measures. In addition, ERP methodology may prove useful in narrowing the range of candidate variables that must be analyzed in order to understand the phenomenon associated with improved behavioral performances (Deutsch et al., 2008).

In the last decade, studies have used ERPs to study emotion recognition. There is evidence that ERPs are an important objective measure of auditory emotional prosody differentiation and recognition (Pinheiro et al., 2011). ERP differences across emotions can be found as early as 200 ms after stimulus onset during both visual and auditory emotional processing (Agrawal et al., 2012; Kotz et al., 2006; Schapkin et al., 2000). The traditional ERP methodology reveals the phase-locked neural activities evoked by a particular cognitive process. Time-frequency (TF) analysis, on the other hand, can reveal the non-phase-locked (induced) neural activity that is hidden in standard averaged ERPs (Makeig et al., 2004; Tallon-Baudry et al., 1999). While ERPs and time-frequency analysis both provide a view of the events in brain's information processing stream, an advantage provided by TF analysis, relative to ERPs, is its potential to understand the brain's parallel processing of information. The oscillations at various frequencies reflect the interaction of multiple neural activities in the service of integrative and dynamically adaptive information processing (Roach and Mathalon, 2008).

This additional benefit of TF analysis, relative to ERPs, may also be manifested in greater sensitivity to the neuronal processes underlying prosody perception in CI users. Additionally, non-phase locked brain activity in the gamma-band range should indicate facilitated prosody recognition modulated by strategy (Hannemann et al., 2007; Lenz et al., 2007). However, although frequency-domain (spectral) analysis has been applied by various researchers in memory and speech perception paradigms (Fuentemilla et al., 2006; Muller et al., 2009), it is yet to be used in emotional prosody analysis in CI users.

The present study aimed to elucidate differences between the effects of the ACE and MP3000 coding strategies on emotional prosody recognition. Based on differences between the two speech coding strategies and the advantage of MP3000 strategy in improving spatial recognition cues, we hypothesized that MP3000 users might perform better compared with the ACE users on prosody identification, which should be reflected in behavioral measures and auditory ERPs. We further hypothesized that induced brain activity obtained with TF analysis should yield complementary effects not observable in the classic ERPs reported above.

In the current study, prosodic information was presented by using neutral, angry and happy tones of voice. Differences were assessed in the behavioral responses, the ERPs as well as TF analysis.

## 2. Materials and method

### 2.1. Participants

Forty right-handed native German speakers (22 females, 18 males), aged 25–60 years (mean = 41.5 years, SD = 7) participated in the experiment. The first group of participants (Group I) consisted of 20 CI users (mean = 42.1 years, SD = 7.01) wearing a Nucleus cochlear

implant system as depicted in Table 1. Subjects had used their implants continuously for at least 12 months and had at least 20% speech in noise perception scores on the Oldenburg sentence test (Wagner et al., 1999) prior to the study. Furthermore, subjects were divided into two subgroups with the aim of comparing two speech-coding strategies (ACE vs. MP3000). The first group of participants (Group IA) consisted of ten individuals (mean age of 42.1, SD = 8.2) with an ACE strategy as their default while their speech processors were programmed with MP3000 strategy for the experiment. Similarly, the remaining ten participants in (Group IB) were CI users (mean = 41.1 SD = 7.3) with MP3000 as their default and ACE as the experimental strategy.

A control group (Group II) comprised age and gender-matched NH participants (age range: 25–55 years; mean = 41 years, SD = 7.1). All participants had normal intelligence and reported no history of psychological or neurological problems. In order to test for depression, Beck's Depression Inventory (BDI) was used (Beck et al., 1996). None of the subjects had clinically-relevant symptoms of a depressive episode. The study was carried out in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the Hannover Medical School. All participants gave written informed consent prior to the recording and received monetary compensation for their participation.

### 2.2. Stimuli

The stimulus material consisted of 150 semantically-neutral German sentences with neutral, happy and angry prosody (50 each) spoken by a trained female German speaker. Stimuli were recorded with a sampling rate of 44.1-kHz and a 16-bit digitizer (Kotz and Paulmann, 2007; Kotz et al., 2003; Wittfoth et al., 2010). All sentences started with personal pronouns (For example, "Sie hat die Zeitung gelesen"; "She has read the newspaper"). The stimulus material was prosodically analyzed using Praat 5.1.19 (Boersma and Weenink, 2005). Table 2, displays the differences in the fundamental frequency (F0), intensity, and duration of the sentences extracted. It was observed that there were significant pitch differences across emotions but no significant intensity differences seen. These significant differences in pitch contours are illustrated in Fig. 1.

### 2.3. Procedure

Testing was carried out in a sound-proofed chamber. Subjects were seated in a comfortable armchair facing a computer monitor placed at a distance of one meter. The stimuli were programmed

**Table 1**  
Demographic data of CI users.

Patient	Age	Gender	Duration of implant-use (in years)	Strategy
1	39	M	3	ACE
2	51	F	3	ACE
3	35	F	3	ACE
4	48	M	4	ACE
5	42	F	3	ACE
6	49	M	2	ACE
7	47	F	2	ACE
8	29	F	3	ACE
9	52	M	4	ACE
10	56	M	5	ACE
11	35	F	3	MP3000
12	43	F	4	MP3000
13	22	F	2	MP3000
14	42	M	2	MP3000
15	39	F	3	MP3000
16	46	F	4	MP3000
17	42	F	3	MP3000
18	48	M	3	MP3000
19	40	M	4	MP3000
20	41	F	3	MP3000

**Table 2**

Averaged acoustical parameters of the presented sentence categories (standard deviations in parenthesis).

	Stimulus	Mean duration (s)	Mean F0 (Hz)	Mean intensity (dB)
Original	Neutral	1.60 (0.3)	157.0 (23.0)	68.6 (1.0)
	Angry	1.70 (0.3)	191.5 (25.0)	70.0 (0.9)
	Happy	1.80 (0.4)	226.6 (24.6)*	67.3 (0.9)

\* Indicate the significant difference in mean F0 for happy prosody vs angry as well as neutral ( $p < .05$ ).

with Presentation Version 14.1 (Neurobehavioral systems) and were presented in a fully-random order via loudspeakers positioned to the left and right of the monitor at a listening level which participants indicated to be sufficiently comprehensible. The task was a three-choice measure with each of the three emotions corresponding to one of the response keys on the response box. Stimuli were presented at a fixed presentation rate with an inter-trial interval of 2500 ms. Participants were asked to decide as accurately as possible whether the presented sentence was spoken in neutral, happy, or angry prosody. The matching of buttons to responses was counterbalanced across subjects within each response group. The experiment included two randomized runs of approximately 13 min each. CI users had their speech processors programmed for ACE and MP3000 strategies. The MP3000 map was optimized to ensure that the overall loudness was similar to that of the ACE map. The study was cross-over in design, hence, for CI users of group (IA), run A involved the use of their conventional speech coding strategy (ACE), whereas in the second run, run B, they used the experimental strategy (MP3000). Similarly for CI users of group (IB), run A entailed the use of their conventional strategy and run B the experimental. To control for familiarity effects the runs were counterbalanced across subjects. Only the responses given after the completion of a sentence were included in later analyses.

#### 2.4. ERP procedure and analysis

The EEG signals were acquired using Brain Vision amplifier system (BrainProducts, Germany, [www.brainproducts.de](http://www.brainproducts.de)). Thirty electrodes were attached to an elastic cap at 10–20 positions (Jasper and Radmussen, 1958) and were referenced to the tip of the nose. In order to control for horizontal and vertical eye movements, a unipolar electro-oculogram was recorded using two electrodes: one placed at the outer canthus, and one below, the right eye. The impedances of the active electrodes were kept below 10 k $\Omega$ . EEG and electro-

oculograms were analog filtered (0.1–100 Hz), and recorded with a sampling rate of 250 Hz. The EEG was recorded continuously on-line and stored for off-line processing.

#### 2.5. Data processing

##### 2.5.1. Preprocessing and artifact rejection

The recorded brain activity was analyzed offline using EEGLAB (Delorme and Makeig, 2004) software (version 9\_0\_4\_5s) running under the MATLAB environment. The data were band-pass filtered from 1 to 30 Hz for ERP peak amplitude computations and 1 to 100 Hz for TF analysis.

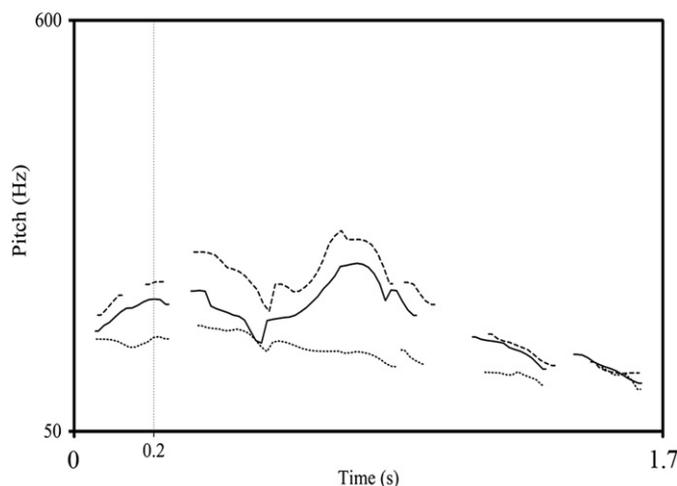
Trials with non-stereotypical artifacts exceeding three standard deviations of an inbuilt probability function (jointprob.m) were removed. ICA was performed with the Infomax ICA algorithm on the continuous data (Debener et al., 2010) with the assumption that the recorded activity is a linear sum of independent components (ICs) arising from brain and non-brain artifact sources. For systematic removal of components representing ocular and cardiac artifacts the EEGLAB plug-in CORRMAP (Viola et al., 2009), enabling semi-automatic component identification, was used. After artifact attenuation by back-projection of all but the artifactual ICs, the clean data were inspected for CI-related artifacts. Furthermore, ICA topographies representing CI artifacts were identified by the centroid on the side of the implanted device and by the cochlear implant pedestal in the time course of the respective component, and were removed. For the CI users, two to three electrodes were removed to avoid direct contact with the CI electromagnetic coil attached to the mastoid. These missing channels were spherically interpolated (Agrawal et al., 2012; Timm et al., 2012). After preprocessing the mean number of artifact-free data epochs that were eventually available across subjects did not differ significantly between emotions.

##### 2.5.2. Peak analysis

The cleaned data were selectively averaged for each emotion condition from the onset of the stimulus, including a 200 ms pre-stimulus baseline, with an 1800 ms ERP time window. The auditory N1 was identified as the negative peak between 100 and 200 ms and the P200 component as the positive peak in the 200–300 ms. In all cases, the baseline-to-peak value was taken as the magnitude of the response. Visual inspection of grand-average waveforms showed that the distribution of ERP effects was predominantly fronto-central. Therefore, only Cz was selected for further analysis. Grand averages for each condition were computed by averaging the single-subject ERP averages for each emotion.

##### 2.5.3. Time-frequency analysis

TF analysis of single trial data was performed using EEGLAB with the inbuilt function newtimef. In order to decompose the signal in the time and frequency domain wavelet analysis was used. Epoched data were transformed into two-dimensional TF planes by convolving the waveforms with a Morlet wavelet at a width of 3 cycles for low, increasing to 15 cycles for high frequency. TF analysis was performed according to the procedures described by Makeig et al. (2004). The percentage of power change in each window relative to the power in the baseline (from –300 to 0 ms pre-onset of the stimulus) was calculated. Evoked power was derived from the TF representation of the averaged signal, yielding only signal portions phase-locked to stimulus onset while omitting non-phase-locked activity. To calculate induced power, the wavelet transformation was applied separately on the signals of each single-trial, and the single-trial results were then averaged. The TF analysis spanned the frequency bands of theta (4–8 Hz), alpha (9–15 Hz), beta (16–34 Hz), and gamma (35–60 Hz). In order to examine the spatial distribution in each frequency band, values were averaged between 0 ms and 1200 ms in selected time windows. Based on both visual inspection and analyses of power changes, windows that showed similar patterns of effects were clustered. This resulted in an early (0–400 ms) and a late (600–1200 ms) time window. On the basis of the acoustic



**Fig. 1.** Pitch contours of the three prosodies. Praat generated pitch contours of neutral (dotted line), angry (solid line) and happy prosody (dashed line) for the original (unsimulated) sentence: “Sie hat die Zeitung gelesen”. The maximum difference in pitch across emotions can be seen between 200 and 1200 ms from the starting of the sentence.

analysis of prosodies, differences were found predominantly in the late window. In order to increase statistical power, and in agreement with previous work (e.g. Paulmann et al., 2011), the statistical analysis was limited to the Cz electrode. Obtained power values were subjected to statistical analysis.

## 2.6. Statistical analysis

The statistical analysis of this crossover design is based on the intention-to-treat population including all randomized patients. The statistical evaluation of the accuracy score, reaction time and ERPs was performed (SPSS 14.0, SPSS Inc. Chicago, Illinois, USA) using repeated measures ANOVA with independent variables *Prosody* (neutral, angry, happy) and *Subject group* (NH, CI). For comparisons of strategy, an ANOVA was performed with factors, *Prosody* (neutral, angry, happy), and *Strategy* (ACE, MP3000). To assess the significance of TF results an ANOVA with factors *Prosody* (neutral, angry, happy) and *Subject group* (NH, CI) was performed for the early as well as the late time window. A similar statistical procedure was applied for comparison of strategies with factors *Prosody* (neutral, angry, happy) and *Strategy* (ACE, MP3000). The relationship between behavioral results and ERPs was studied by correlating accuracy rate and P200 amplitude with Pearson product correlation. As the ANOVA results of reaction time and latency differences were not significant across groups and condition, they were excluded from correlation analysis. The correlation analysis was also conducted to study a relationship between accuracy rate and induced gamma power for late time window only. All hypotheses were tested to a significance level of 5% (two-sided). In order to correct for sphericity violation the Greenhouse–Geisser correction was used in relevant cases.

## 3. Results

### 3.1. Accuracy data

The group means of the performance accuracy scores (% correct) are depicted in Fig. 2(A). An ANOVA revealed a significant main effect of group [ $F(1, 38) = 53.32, p = 0.001$ ], showing that performance of NH controls was better than that of CI users. There was no main effect of prosody.

The comparison of performance across strategies revealed a significant main effect of strategy [ $F(1, 38) = 5.156, p = .029$ ], demonstrating that performance accuracy was higher in MP3000 (72  $\pm$  17%) compared to ACE (61  $\pm$  15%). A significant interaction between strategy and prosody was also observed [ $F(1, 38) = 18.659, p = .001$ ]. Follow-up paired t-tests revealed higher accuracy for happy prosody recognition with MP3000 [ $t(19) = 3.164, p = .005$ ] compared to ACE. However, the accuracy scores between the two strategies were comparable for angry and neutral prosody. No other significant differences could be observed.

Note that half of the subjects used ACE as their default strategy and MP3000 as experimental while the other half used MP3000 as their default strategy and ACE as experimental. Our results revealed no significant difference in regard to the accuracy rates [ $p > .050$ ] between the groups, indicating that the subjects with default ACE were comparable to subjects with experimental ACE. Similar results were obtained for MP3000 strategy, suggesting that effects were not biased by duration of use.

### 3.2. Reaction times

The group mean reaction times are depicted in Fig. 2(B). Analysis of the reaction time data revealed a significant main effect of

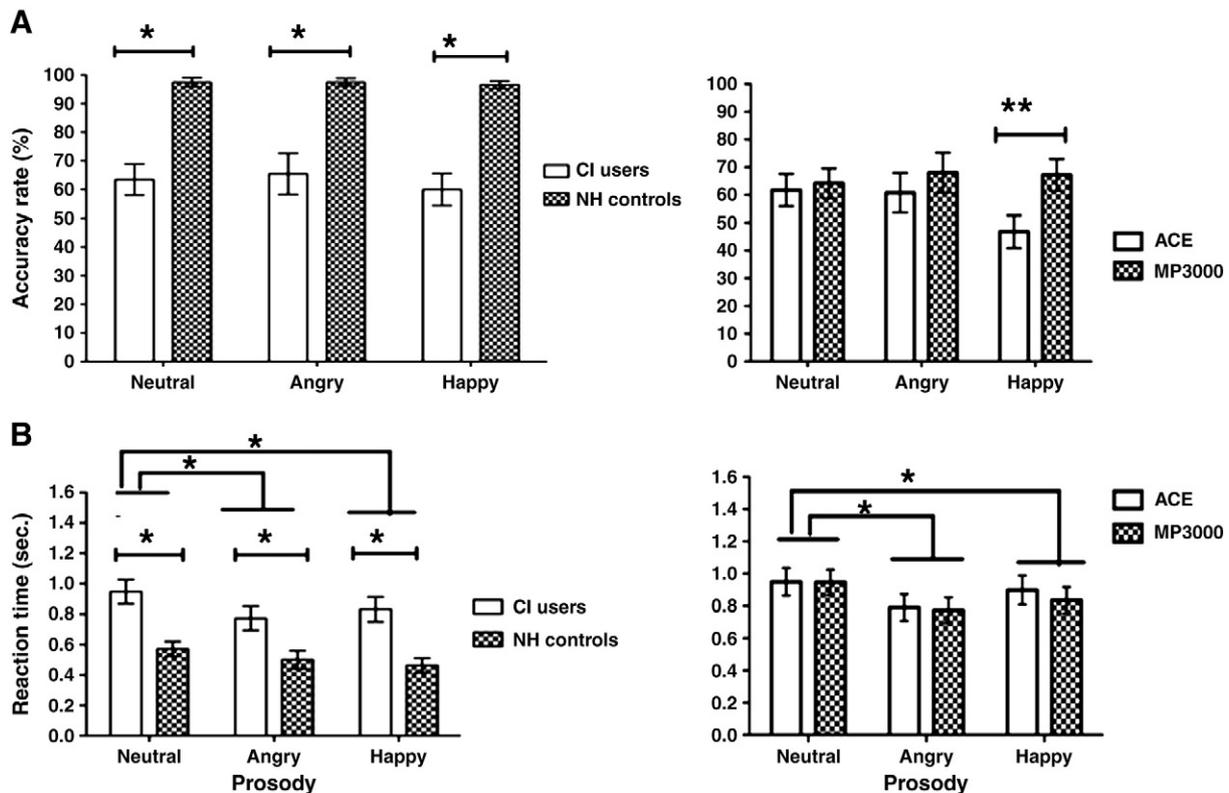


Fig. 2. Accuracy rate and reaction time. Top (A): accuracy rate for NH controls vs. CI, ACE users vs. MP3000 users for neutral, angry and happy emotional prosody recognition in percentage. Bottom (B): post offset reaction time for emotional prosody recognition in seconds. The significance is indicated with asterisk with  $p < 0.05$ .

group [ $F(2,18) = 9.090, p = .001$ ], showing that reaction time of CI users (840 ms) was longer compared to that of NH controls (520 ms) by an average of 320 ms (collapsed across prosody). Furthermore, a significant main effect of prosody [ $F(2,38) = 14.337, p = .001$ ] was found, although there was no interaction. Breakdown analysis of the main effects revealed that reaction times in recognition of happy (480 ms), [ $t(39) = 3.418, p = .020$ ] and angry (500 ms), [ $t(39) = 3.536, p = .022$ ] prosody were significantly shorter than those of neutral (600 ms) for CI users. No significant difference was found between happy and angry prosody.

Similarly, on prosody recognition in CI users, a significant main effect of prosody was observed [ $F(2,38) = 16.315, p = .001$ ]. Breakdown analysis showed that subjects took less time to recognize happy (820 ms) [ $t(39) = 5.081, p = .001$ ] and angry (810 ms) [ $t(39) = 2.672, p = .011$ ] prosody compared to neutral (910 ms) (collapsed across strategies). Data yielded no significant difference in reaction times between happy and angry prosody. There was no significant interaction between strategy used and prosody recognition.

The same procedure described above was used to check if there was an effect of default vs. experimental strategy use. Analysis revealed no significant differences between the two [ $p > .05$ ] either for ACE or for MP3000.

### 3.3. Event-related potentials

Fig. 3 depicts the ERP waveforms for the different subject groups across three emotional prosodies. Mean latencies and amplitudes for N100 and P200 peaks are presented in Table 3.

#### 3.3.1. N100

An ANOVA on N100 latency revealed significant differences between the NH and CI users [ $F(1,38) = 6.080, p = 0.002$ ] indicating longer N100 latencies for CI users. We found no significant main effect of prosody and no interaction between prosody and subject group. The comparison of strategy revealed no significant effect of prosody. There was no significant interaction between strategy and prosody.

For the amplitude analysis, ANOVA revealed significantly greater N100 amplitude for NH controls compared with CI users [ $F(1,38) = 6.378, p = 0.003$ ]. However, the interaction between prosody and group did not reach significance. The comparison of strategies showed that there was no significant main effect of prosody or strategy, and no interaction.

#### 3.3.2. P200

A significant main effect of subject group [ $F(1,38) = 20.907, p = 0.001$ ] was observed. Pair-wise comparisons revealed that P200 latency was shorter in NH controls compared with CI users by an average of 35 ms (collapsed across prosodies). No significant main effect of prosody was observed, and there was no group-by-prosody interaction. The comparison of strategies revealed no significant effects on P200 latency measures.

A significant group effect [ $F(1,38) = 28.245, p = 0.001$ ] indicated a reduced P200 peak amplitude in CI users compared to NH controls (by an average of 3.8  $\mu\text{V}$  collapsed across prosodies). On comparison of strategies, a significant effect of strategy use [ $F(1,18) = 12.395, p = 0.006$ ] was observed, indicating that the P200 peak amplitude was significantly larger in MP3000 users compared to ACE. Follow up t-tests revealed that MP3000 strategy use yielded larger amplitude for happy prosody [ $t(19) = 4.240, p = .001$ ] compared with ACE, but not for angry and neutral prosody.

### 3.4. Time-frequency analysis

Mean evoked and induced powers were calculated in two time windows (0–400 ms) and (600–1200 ms) for four frequency bands: theta

(4–8 Hz), alpha (9–15 Hz), beta (16–34 Hz), and gamma (35–60 Hz). There were no significant differences in either alpha or beta activity.

#### 3.4.1. Theta

Baseline-corrected spectral theta power evoked by prosody showed a clear peak in the early and late time windows. Statistical comparison of differences across groups revealed a significant main effect of group in the early window [ $F(1,19) = 9.779, p = 0.003$ ]. Theta activity was larger in NH controls compared to CI users. However, the main effect of prosody did not reach significance.

The comparison of theta power across strategies revealed no significant differences in the early time window. In the late window, a significant main effect of strategy was observed [ $F(1,18) = 10.02, p = 0.005$ ], reflecting higher theta power when participants used the MP3000 strategy. However, the main effect of prosody did not reach significance.

#### 3.4.2. Gamma

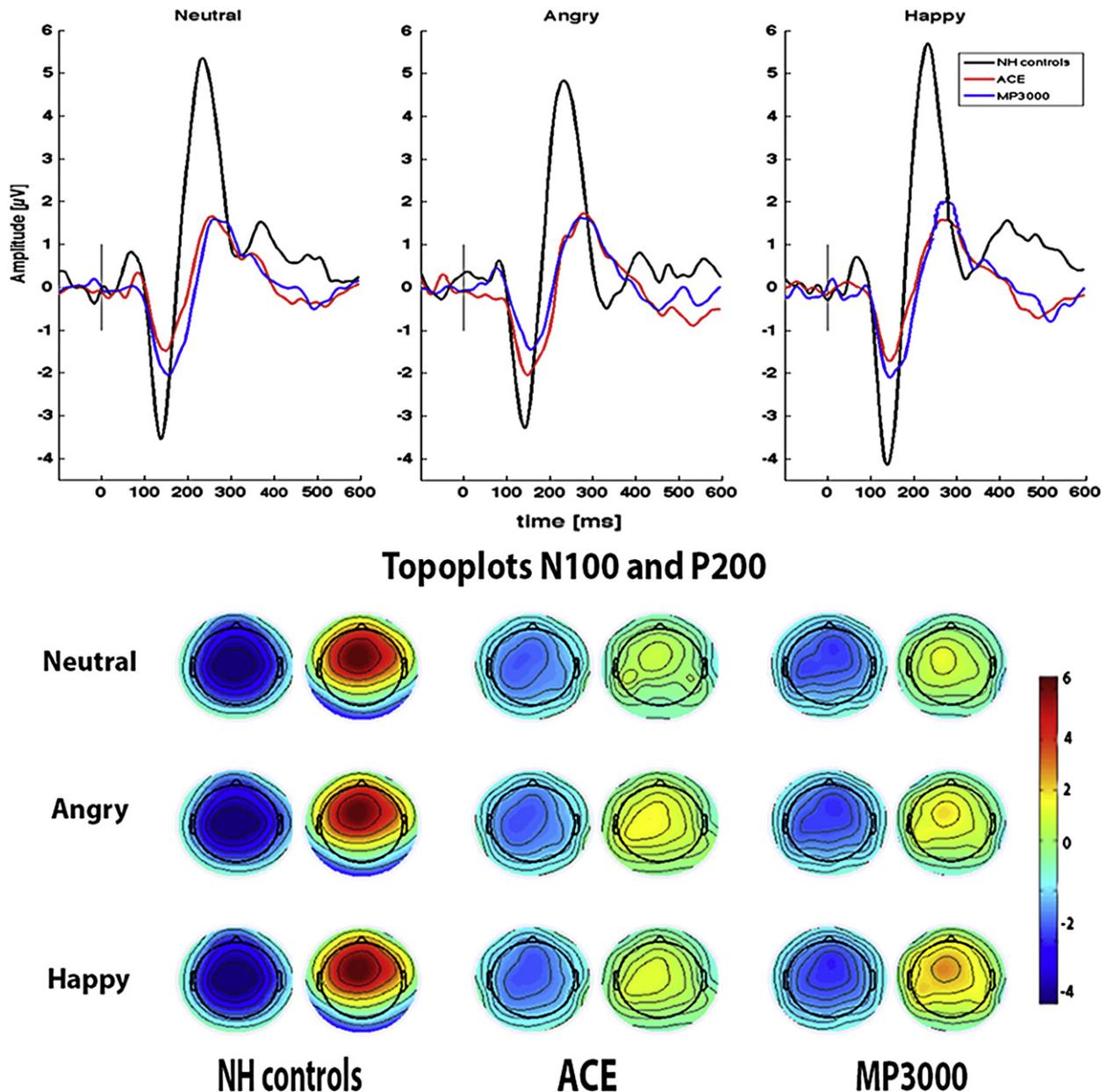
Fig. 4 depicts the baseline-corrected TF plots for induced gamma activity at Cz for CI users. Recurring patches of gamma enhancement were observable in the TF plots. Bursts of gamma activity after stimulus onset were found for induced gamma activity. ANOVA revealed that in NH controls induced gamma activity was significantly larger compared with CI users irrespective of the prosody studied for the early [ $F(1,18) = 6.20, p = 0.005$ ] as well as the late time window, [ $F(1,18) = 5.312, p = 0.010$ ]. No interaction was observed between prosody and groups.

On comparison of induced gamma activity across strategies in the early time window, a significant main effect of strategy [ $F(1,38) = 8.172, p = 0.020$ ] was observed. In this window induced gamma power showed a prominent peak in a time window from 180 to 250 ms after stimulus onset in MP3000 users, whereas ACE users showed almost no gamma-band change, reflecting lack of prosody recognition. The interaction between prosody and strategy was also significant [ $F(1,38) = 3.779, p = 0.042$ ]. Follow up comparisons indicated that gamma activity induced by happy prosody was significantly stronger for MP3000 users [ $t(19) = 2.789, p = .021$ ] compared to ACE. The neutral and angry prosody did not show any significant difference. Similarly, in the later window, induced gamma activity increased along with the increase in the prosody recognition ability. MP3000 users displayed higher induced gamma power compared to ACE [ $F(1,38) = 8.881, p = 0.020$ ]. This activity peaked at 1050 ms after stimulus onset. A significant interaction of strategy and prosody was observed [ $F(1,38) = 4.241, p = 0.033$ ]. Follow up t-tests revealed that gamma activity induced by happy prosody for MP3000 was stronger than that for ACE, [ $t(19) = 2.430, p = .025$ ], but not with neutral and angry prosody. No other comparisons yielded significant results.

Taken together, the results in both the early and late time windows revealed that ACE and MP3000 users differed significantly, in particular in showing that induced gamma power was more pronounced for MP3000 users compared to ACE users in response to happy prosody.

### 3.5. Correlations

Pearson product-moment correlation was carried out to study the correlation of the P200 peak amplitude with the behavioral performance (accuracy rate only as differences in reaction time across subject groups were not significant). Significant correlations were observed for happy prosody between accuracy rate and P200 amplitude in NH controls [ $r = 0.736; p = 0.019$ ] as well as in MP3000 users [ $r = 0.656; p = 0.036$ ], reflecting that the P200 peak amplitude increased with higher accuracy of happy prosody recognition. No significant correlations were observed for angry or neutral prosody. The correlation analysis of accuracy rate and gamma band power for happy prosody revealed, for



**Fig. 3.** ERP waveforms for three emotional prosodies for NH controls, ACE and MP3000 users. Average ERP waveforms recorded at the Cz electrode for NH controls (black), ACE users (red) and MP3000 users (blue) for all three emotional [neutral, angry and happy] stimuli from 100 ms before onset to 600 ms after the onset of the sentences with respective scalp topographies at N100 and P200 peak (X-axis: latency in milliseconds, Y-axis: amplitude in  $\mu\text{V}$ ). Top left: N100-P200 waveform for neutral, Middle: waveform for angry and right for happy prosody. Bottom left scale topography of NH controls, middle: ACE users, right: MP3000 users for N100 and P200 respectively.

NH controls, a significant positive correlation between gamma band power and the accuracy rate [ $r=0.636$ ;  $p=0.019$ ], indicating that the high induced gamma power was also associated with better happy prosody recognition. Similar observations were made for the MP3000 users [ $r=0.642$ ,  $p=0.036$ ]. The correlation analysis for neutral and angry prosody did not reach significance. The correlation analysis therefore revealed a specific, but clear, link between emotion recognition, P2 peak amplitude, and induced gamma power for happy prosody in MP3000 users and in NH controls, but not in ACE users.

#### 4. Discussion

In this cross-over study significant differences were found in emotion recognition across subjects reflected by both behavioral and electrophysiological measures. Comparisons between CI and NH listeners indicated that CI users had difficulty recognizing prosody. This could

be attributed to the limited dynamic range of an implant by which subjects have limited access to pitch, voice quality, vowel articulation, and spectral envelope cues; features that are thought to be essential for emotional voice recognition. Further, comparison of speech coding strategies revealed that the MP3000 strategy has an advantage over ACE strategy on happy prosody recognition.

##### 4.1. Behavioral findings

In general, all participants took longer to recognize neutral prosodic sentences compared with sentences spoken in angry or happy prosody. A possible explanation might be that, in the emotional judgment of prosody the non-ambiguous emotional associations are readily available, resulting in faster recognition. In contrast, neutral stimuli may elicit positive or negative associations which otherwise

**Table 3**

Peak mean latency and amplitude of the N100 and P200 (standard deviations in parenthesis).

Subjects	Neutral	Angry	Happy
<i>N100 peak latency (ms)</i>			
NH	137 (11.5)	138 (13.5)	140 (10.1)
ACE	145 (22.0)	156 (23.4)	151 (22.0)
MP3000	154 (22.0)	155 (22.2)	154 (21.8)
<i>N100 peak amplitude (<math>\mu</math>V)</i>			
NH	−3.90 (1.8)	−3.90 (1.5)	−4.00 (1.9)
ACE	−2.31 (1.3)	−2.82 (1.3)	−2.60 (1.2)
MP3000	−2.81 (1.8)	−2.50 (1.3)	−2.55 (1.4)
<i>P200 peak latency (ms)</i>			
NH	240 (16.6)	240 (20.2)	234 (10.0)
ACE	259 (25.0)	270 (26.1)	270 (24.1)
MP3000	262 (25.0)	271 (25.6)	271 (25.2)
<i>P200 peak amplitude (<math>\mu</math>V)</i>			
NH	5.90 (1.5)*	6.00 (1.5)*	6.20 (1.8)*
ACE	2.23 (1.2)	2.38 (0.9)	2.34 (0.9)
MP3000	2.21 (1.1)	2.31 (0.8)	2.81 (0.9)**

\* Indicate the significant differences in mean P200 amplitude for NH subjects compared with ACE and MP3000 users.

\*\* Indicates significant P200 amplitude difference between ACE and MP3000 users ( $p < .05$ ).

may not exist (Grandjean et al., 2008). Here, the reaction times simply reflect a longer decision time for neutral prosody.

The happy prosody was recognized with the highest accuracy compared with the other two prosodies. There is evidence to suggest that negative stimuli are less expected and take more effort to process compared with positive stimuli hence happy emotions are more socially salient (Johnstone et al., 2006; Lang and Bradley, 2009). Further, the rate of happiness linearly increased as a function of the pitch variations, suggesting a graded processing of neutral and emotional voice, consistent with other studies (Spreckelmeyer et al., 2009). Thus the high accuracy might be due to the social importance of happy emotion and the additional pitch cues. Moreover, NH subjects listening to CI simulations were significantly less accurate and took longer time to recognize emotional prosodic information compared with original unstimulated stimuli as observed in our recent study (Agrawal et al., 2012).

#### 4.2. ERPs

In line with the behavioral results, CI users exhibited prolonged ERP latencies and reduced amplitudes compared with NH controls. This finding is supported by previous research (Luo and Fu, 2007), which has reported that the amount of sensory information processed through a CI is less compared to an intact cochlea resulting in reduced synchronization of neuronal activity required in generating auditory evoked potentials (Groenen et al., 2001; Sandmann et al., 2009). Although CI users had prolonged latency and reduced amplitude ERPs, the structure of these potentials was similar to those recorded from NH controls. This implies that despite the limited input provided by CIs, the central auditory system processes the prosodic stimuli consistently in a relatively normal fashion (Koelsch et al., 2004). The ERP data for emotional prosody perception recorded in all the participants' demonstrated differential electrophysiological responses in sensory-perceptual component of emotional relative to neutral prosody. The auditory N100 component is a marker of physical characteristics of stimuli such as temporal pitch extraction (Seither-Preisler et al., 2006). Evidence exists in the literature advocating the N100 as the first stage of emotional prosody processing (Schirmer and Kotz, 2003). In the current study, N100 amplitude was not significantly different across strategies, suggesting that early stages of prosody recognition might be adversely affected in some way by stimulus characteristics. However, the N100 is modulated by various factors including attention, motivation, arousal,

fatigue, complexity of the stimuli, and methods of recording etc. (Pinheiro et al., 2011). Thus, it is not possible to delineate the reasons for the presence of the N100 since one cannot rule out the contribution of above-mentioned factors to the observed results.

The P200 component reflects the initial encoding of the emotion (Balconi and Pozzoli, 2003). Previous studies have reported that emotional stimuli elicit larger ERP waveforms than neutral stimuli, often as early as 200 ms after stimulus onset (Vanderploeg et al., 1987). This initial emotional encoding seems to be influenced by stimulus pitch (Pantev et al., 1996). The lack of a statistical difference between the intensity of sentences spoken with angry prosody versus spoken with happy indicates that the differences observed in P200 amplitudes were not simply due to loudness or emphasis. Thus, it seems unlikely that an attention-related effect due to an unexpected change of loudness could account for the observed ERP differences across emotional prosodies (Table 2 acoustic parameters). More importantly, a significant positive linear correlation between P200 amplitude and rate of accuracy for happiness recognition was observed. This result is consistent with the observation that the brain response (P200) is sensitive to the stimulus pitch modulation (Pantev et al., 1996). Thus, the brain response observed in the present study was sensitive to the pitch variations of emotional prosody rather than loudness.

Furthermore, the amplitude of the P200 component in MP3000 users was larger compared to ACE. This could be due to the fact that the MP3000 strategy avoids repetitive stimulation of neuronal ensembles (Buechner et al., 2008), by selecting components that are more widely dispersed across the spectrum to avoid clustering of stimulated channels (Nogueira et al., 2005). Thus, bands selected by the psychoacoustic model in MP3000 extract the most meaningful components of the audio signal based on normal cochlear physiology. Due to this, only the relevant components are transferred and the redundant components are masked. Since this strategy is based on a normal-hearing principle, the extraction of fundamental (F0) frequency cues fares better compared with ACE, resulting in an improved recognition of happy prosody in the present study.

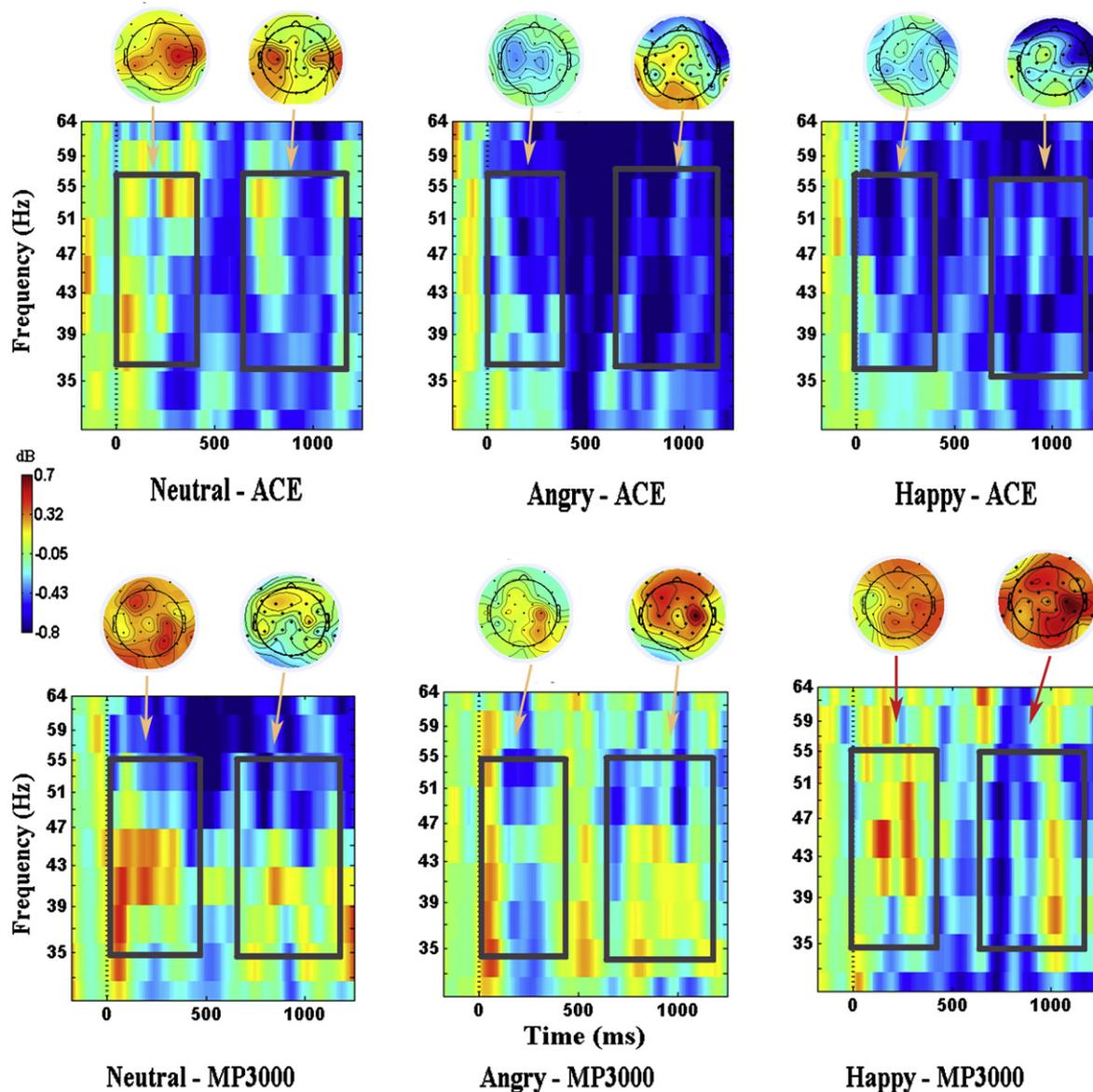
#### 4.3. Time-frequency results

In accordance with conducting conventional analysis on ERP waveforms, we also used time-frequency analysis. The focus of time-frequency analysis is on oscillations, which are readily detected in the EEG when a frequency decomposition is performed. To our knowledge, our study is the first attempt to show that the processing of emotional prosody in CI users can be distinguished in both early and late stages of brain oscillations during prosody recognition. Our results support the view that frequency-specific EEG responses differ from each other systematically as a function of the stimulus type (i.e. emotional vs. neutral). Theta and gamma bands showed a significant power increase with the emotional stimuli, whereas alpha and beta frequencies were not modulated by the prosody; an observation which is in line with previously reported effects (Aftanas et al., 2001, 2004; Knyazev et al., 2009).

Overall spectral power was larger for NH controls compared to CI users as evident from ERP results. Consistent with the idea that NH subjects have a better hearing resolution than CI users and that stimulus presentation is well received. On the other hand the larger gamma band power in CI users could be indicative of additional neural activation in CI users to compensate for less efficient processing during extraction of stimulus features and integration of the perceived sensory input. Imaging studies support this assumption by demonstrating that task-related compensatory activations are higher in response to developmental deviations compared to normal maturational processes (Durstun et al., 2003; Sheridan and Hausdorff, 2007).

In regard to the comparisons of CI strategies used in the present experiment, MP3000 users showed significantly larger power at theta frequencies compared to ACE users. These findings are in line with the

## Gamma Power (-200 - 1200 ms)



**Fig. 4.** Induce gamma power plots for ACE and MP3000 users. The average gamma power computed at Cz electrode for ACE users and MP3000 users for all three emotional prosodies (neutral, angry and happy) for (–200–1500) time window in range of 25 to 70 Hz. Highlighted area along with scalp plots represent induced gamma power in early (0–400 ms) and in late (600–1200 ms) time windows at significance ( $p < 0.01$ ).

literature, where it has been observed that there is a general tendency for higher theta along with an increased motivation and emotional significance of the stimuli (Aftanas et al., 2001). This clearly explains the higher theta power for emotional stimuli, particularly happy emotion in MP3000 compared to ACE users, suggesting that slow cerebral oscillations are suitable to study the processes related to emotion recognition. Observed differences can be attributed to the psychoacoustic model in MP3000 strategy. This algorithm increases the dynamic range, thereby improving the pitch as well as finer feature recognition.

Two typical peaks in gamma activity were observed for prosody recognition in both the early and the late time windows. The early gamma activity occurred at about 100–250 ms after stimulus onset (induced), while the late peak commenced after 600 ms post-stimulus (induced). In the present study, MP3000 users showed significantly higher power at gamma-band frequencies for the recognition of emotions compared with ACE users. MP3000 users showed the highest power for happy emotion recognition and the lowest power for neutral stimuli.

In contrast ACE users displayed higher power for neutral compared to happy prosody. There is strong evidence that gamma power in the initial 250 ms is a correlate of sensory processing (Busch et al., 2004; Karakas and Basar, 1998) reflecting bottom-up processes driven by stimulus features such as loudness (Schadow et al., 2007) and pitch (Sedley et al., 2012). There is also evidence that early gamma activity represents an interface of bottom-up factors and top-down processes (Busch et al., 2006). Thus, it implies that in the current study, the MP3000 strategy encoded the physical characteristics better than ACE, which in turn improved the prosody recognition especially for happy prosody that had maximum pitch modulation. Similarly, increased induced gamma band power was also observed in the late time window of 600 to 1200 ms peaking at 1050 ms. It seems likely that the second burst of long-latency gamma synchrony reported in the late time window may be related to the maintenance or refinement of the attention network established by the first burst. Thus, the selection of certain neuronal groups for integration into a

large-scale synchronous gamma network accounts for the increased integration of attended information.

Similarly, in the late time window the variation of acoustic features was at its peak. This suggests that these features were well received and coded by the brain to differentiate emotions effectively. MP3000 users had the highest gamma power during happy prosody recognition, whereas ACE users had the lowest power. Since previous studies have shown a similar gamma-band power increase locked to visual stimuli conditioned to emotional pictures (Stolarova et al., 2006), the current results support a specific role of gamma activity (35–60 Hz) for emotionally triggered functional activation states, seemingly irrespective of the sensory modality (Stolarova et al., 2006). High-frequency components, such as gamma-band activity, have been associated with ‘local’ computations and the binding of fine-structured information (Tallon-Baudry et al., 1999), whereas low frequency components such as theta involve more ‘global’ computations and are possibly amodal (Yordanova et al., 2002). Our results suggest that a more analytic (‘local’) processing mode was engaged in the emotion identification with MP3000, compared to the ACE. The ‘burst of gamma’ observed in perceptual testing originates from a transient synchrony of neural populations involved in the processing of a sensory event, i.e. the stimulus enters the realm of conscious recognition (Hopfield and Brody, 2000; Mukamel et al., 2005). The positive correlation between gamma activity and neuronal spiking is a robust finding reported in aforementioned studies. Thus, the MP3000 strategy leads to higher neuronal firing rate compared to ACE which might have improved the identification of the prosody with this strategy. Gamma band responses also show a direct relationship with task demands (Senkowski and Herrmann, 2002; Yordanova et al., 1997). Hence, it seems reasonable to assume higher activation of processing resources in MP3000 users for task-relevant stimuli, reflected by enhanced activity in the gamma-range. Furthermore, the high positive correlation between accuracy rate and gamma band power reflected an advantage of MP3000 over ACE on prosody recognition. These findings are in line with previous reports in which an association between power in gamma frequencies and response accuracy was more pronounced in good performers compared with bad performers. This difference was particularly manifested around the onset of the test stimuli (Kaiser et al., 2008). Thus the results suggest the relevance of gamma power for optimal differentiation between stimulus characteristics. However, there are some limitations that should be taken into consideration while making these conclusions. First, only prosodies recorded by one speaker were used, which constrains the generalization of the findings. Second, only three emotions were intoned via the sentences. Despite these limitations the results are very promising and reflect the advantage of using ERP and TF analysis in evaluating CI users’ abilities that may not be evident in behavioral measures.

Taken together, we have shown that subjects using the ACE strategy had more difficulty perceiving emotional stimuli compared with MP3000 users. These difficulties were reflected in behavioral scores, ERPs, and TF measures. It was observed that subjects with ACE had to be more attentive to distinguish the emotions compared with NH and MP3000 users. High-frequency oscillations typically show lower amplitudes compared to low-frequency oscillations and therefore are not visible in ERPs and would not have contributed to the peak amplitudes (Edwards et al., 2005). Thus, these findings lend strong support for the hypothesis that the MP3000 strategy has an advantage over ACE on happy prosody perception and that this can be confirmed with behavioral and EEG findings. In addition, gamma band activity may serve as a relevant measure for comparing performance of speech coding strategy for prosody recognition in CI users when ERPs are not sufficiently sensitive. Our findings of the P200 amplitude differences corroborate the current literature on the use of auditory ERPs as an objective measure in CI research (Sandmann et al., 2009; Viola et al., 2011). Here, we showed that even finer acoustic differences might be appropriate to compare hearing abilities of CI speech coding strategies.

## Funding information

This research was supported by the grants from the Georg Christoph Lichtenberg Stipendium of Lower-Saxony, Germany, and was partially supported by the Fundacao para a Ciencia e Tecnologia, Lisbon, Portugal (SFRH/BD/37662/2007) to F.C.V.

## Acknowledgments

We thank all participants for their support and their willingness to be part of this study.

## References

- Aftanas, L., Varlamov, A., Pavlov, S., Makhnev, V., Reva, N., 2001. Event-related synchronization and desynchronization during affective processing: emergence of valence-related time-dependent hemispheric asymmetries in theta and upper alpha band. *The International Journal of Neuroscience* 110 (3–4), 197–219.
- Aftanas, L.L., Reva, N.V., Savotina, L.N., Makhnev, V.P., 2004. Neurophysiological correlates of induced discrete emotions in humans: an individual analysis. *Rossiiskii Fiziologicheskii Zhurnal Imeni I.M. Sechenova* 90 (12), 1457–1471.
- Agrawal, D., Timm, L., Viola, F.C., Debener, S., Buechner, A., Dengler, R., et al., 2012. ERP evidence for the recognition of emotional prosody through simulated cochlear implant strategies. *BMC Neuroscience* 13 (1), 113.
- Balconi, M., Pozzoli, U., 2003. Face-selective processing and the effect of pleasant and unpleasant emotional expressions on ERP correlates. *International Journal of Psychophysiology* 49 (1), 67–74.
- Beck, A.T., Steer, R.A., Ball, R., Ranieri, W., 1996. Comparison of Beck Depression Inventories-IA and -II in psychiatric outpatients. *Journal of Personality Assessment* 67 (3), 588–597.
- Boersma, P., Weenink, D., 2005. Praat: doing phonetics by computer.
- Buechner, A., Brendel, M., Krueger, B., Frohne-Buchner, C., Nogueira, W., Edler, B., et al., 2008. Current steering and results from novel speech coding strategies. *Otology & Neurology* 29 (2), 203–207.
- Busch, N.A., Debener, S., Kranczioch, C., Engel, A.K., Herrmann, C.S., 2004. Size matters: effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clinical Neurophysiology* 115 (8), 1810–1820.
- Busch, N.A., Schadow, J., Frund, I., Herrmann, C.S., 2006. Time-frequency analysis of target detection reveals an early interface between bottom-up and top-down processes in the gamma-band. *NeuroImage* 29 (4), 1106–1116.
- Debener, S., Thorne, J., Schneider, T.R., Viola, F.C., 2010. Using ICA for the analysis of multi-channel EEG data. In: Debener, M.U.S. (Ed.), *Simultaneous EEG and fMRI*. Oxford University Press, New York, NY, pp. 121–135.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* 134 (1), 9–21.
- Deutsch, C.K., Dube, W.V., McIlvane, W.J., 2008. Attention deficits, Attention-Deficit Hyperactivity Disorder, and intellectual disabilities. *Developmental Disabilities Research Reviews* 14 (4), 285–292.
- Durston, S., Davidson, M.C., Thomas, K.M., Worden, M.S., Tottenham, N., Martinez, A., et al., 2003. Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *NeuroImage* 20 (4), 2135–2141.
- Edwards, E., Soltani, M., Deouell, L.Y., Berger, M.S., Knight, R.T., 2005. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. *Journal of Neurophysiology* 94 (6), 4269–4280.
- Fu, Q.J., Chinchilla, S., Nogaki, G., Galvin III, J.J., 2005. Voice gender identification by cochlear implant users: the role of spectral and temporal resolution. *The Journal of the Acoustical Society of America* 118 (3 Pt 1), 1711–1718.
- Fuentemilla, L., Marco-Pallares, J., Grau, C., 2006. Modulation of spectral power and of phase resetting of EEG contributes differentially to the generation of auditory event-related potentials. *NeuroImage* 30 (3), 909–916.
- Grandjean, D., Sander, D., Scherer, K.R., 2008. Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and Cognition* 17 (2), 484–495.
- Groenen, P.A., Beynon, A.J., Snik, A.F., van den Broek, P., 2001. Speech-evoked cortical potentials and speech recognition in cochlear implant users. *Scandinavian Audiology* 30 (1), 31–40.
- Hannemann, R., Obleser, J., Eulitz, C., 2007. Top-down knowledge supports the retrieval of lexical information from degraded speech. *Brain Research* 1153, 134–143.
- Hopfield, J.J., Brody, C.D., 2000. What is a moment? “Cortical” sensory integration over a brief interval. *Proceedings of the National Academy of Sciences of the United States of America* 97 (25), 13919–13924.
- Hopyan, T., Gordon, K.A., Papsin, B.C., 2011. Identifying emotions in music through electrical hearing in deaf children using cochlear implants. *Cochlear Implants International* 12 (1), 21–26.
- House, W.F., 1994. Cochlear implants: it's time to rethink. *The American Journal of Otology* 15 (5), 573–587.
- Jasper, H.H., Radmussen, T., 1958. Studies of clinical and electrical responses to deep temporal stimulation in men with some considerations of functional anatomy. *Research Publications - Association for Research in Nervous and Mental Disease* 36, 316–334.

- Johnstone, T., van Reekum, C.M., Oakes, T.R., Davidson, R.J., 2006. The voice of emotion: an fMRI study of neural responses to angry and happy vocal expressions. *Social Cognitive and Affective Neuroscience* 1 (3), 242–249.
- Kaiser, J., Heidegger, T., Lutzenberger, W., 2008. Behavioral relevance of gamma-band activity for short-term memory-based auditory decision-making. *The European Journal of Neuroscience* 27 (12), 3322–3328.
- Karakas, S., Basar, E., 1998. Early gamma response is sensory in origin: a conclusion based on cross-comparison of results from multiple experimental paradigms. *International Journal of Psychophysiology* 31 (1), 13–31.
- Knyazev, G.G., Slobodskoj-Plusnin, J.Y., Bocharov, A.V., 2009. Event-related delta and theta synchronization during explicit and implicit emotion processing. *Neuroscience* 164 (4), 1588–1600.
- Koelsch, S., Wittfoth, M., Wolf, A., Muller, J., Hahne, A., 2004. Music perception in cochlear implant users: an event-related potential study. *Clinical Neurophysiology* 115 (4), 966–972.
- Kotz, S.A., Paulmann, S., 2007. When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research* 1151, 107–118.
- Kotz, S.A., Meyer, M., Alter, K., Besson, M., von Cramon, D.Y., Friederici, A.D., 2003. On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and Language* 86 (3), 366–376.
- Kotz, S.A., Meyer, M., Paulmann, S., 2006. Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design. *Progress in Brain Research* 156, 285–294.
- Lang, P.J., Bradley, M.M., 2009. Emotion and the motivational brain. *Biological Psychology* 84 (3), 437–450.
- Lenz, D., Schadow, J., Thaerig, S., Busch, N.A., Herrmann, C.S., 2007. What's that sound? Matches with auditory long-term memory induce gamma activity in human EEG. *International Journal of Psychophysiology* 64 (1), 31–38.
- Loizou, P.C., 1999. Signal-processing techniques for cochlear implants. *IEEE Engineering in Medicine and Biology Magazine* 18 (3), 34–46.
- Luo, X., Fu, Q.J., 2007. Frequency modulation detection with simultaneous amplitude modulation by cochlear implant users. *The Journal of the Acoustical Society of America* 122 (2), 1046–1054.
- Makeig, S., Debener, S., Onton, J., Delorme, A., 2004. Mining event-related brain dynamics. *Trends in Cognitive Sciences* 8 (5), 204–210.
- Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., Malach, R., 2005. Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. *Science* 309 (5736), 951–954.
- Muller, V., Gruber, W., Klimesch, W., Lindenberger, U., 2009. Lifespan differences in cortical dynamics of auditory perception. *Developmental Science* 12 (6), 839–853.
- Nogueira, W., Buechner, A., Lenarz, T., Edler, B., 2005. A psychoacoustic “NofM”-type speech coding strategy for cochlear implants. *Journal on Applied Signal Processing, Special Issue on DSP in Hearing Aids and Cochlear Implants, Eurasip*. 127 (18), 3044–3059.
- Pantev, C., Elbert, T., Ross, B., Eulitz, C., Terhardt, E., 1996. Binaural fusion and the representation of virtual pitch in the human auditory cortex. *Hearing Research* 100 (1–2), 164–170.
- Paulmann, S., Ott, D.V., Kotz, S.A., 2011. Emotional speech perception unfolding in time: the role of the basal ganglia. *PLoS One* 6 (3), e17694.
- Pinheiro, A.P., Galdo-Alvarez, S., Rauber, A., Sampaio, A., Niznikiewicz, M., Goncalves, O.F., 2011. Abnormal processing of emotional prosody in Williams syndrome: an event-related potentials study. *Research in Developmental Disabilities* 32 (1), 133–147.
- Roach, B.J., Mathalon, D.H., 2008. Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin* 34 (5), 907–926.
- Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jancke, L., Dillier, N., et al., 2009. Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain* 132 (Pt 7), 1967–1979.
- Schadow, J., Lenz, D., Thaerig, S., Busch, N.A., Frund, I., Rieger, J.W., et al., 2007. Stimulus intensity affects early sensory processing: visual contrast modulates evoked gamma-band activity in human EEG. *International Journal of Psychophysiology* 66 (1), 28–36.
- Schapkin, S.A., Gusev, A.N., Kuhl, J., 2000. Categorization of unilaterally presented emotional words: an ERP analysis. *Acta Neurobiologiae Experimentalis* 60 (1), 17–28.
- Scherer, K.R., 2003. Vocal communication of emotion: a review of research paradigms. *Speech Communication* 40, 227–256.
- Schirmer, A., Kotz, S.A., 2003. ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience* 15 (8), 1135–1148.
- Sedley, W., Teki, S., Kumar, S., Overath, T., Barnes, G.R., Griffiths, T.D., 2012. Gamma band pitch responses in human auditory cortex measured with magnetoencephalography. *NeuroImage* 59 (2), 1904–1911.
- Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., Lutkenhoner, B., 2006. Evidence of pitch processing in the N100m component of the auditory evoked field. *Hearing Research* 213 (1–2), 88–98.
- Senkowski, D., Herrmann, C.S., 2002. Effects of task difficulty on evoked gamma activity and ERPs in a visual discrimination task. *Clinical Neurophysiology* 113 (11), 1742–1753.
- Sheridan, P.L., Hausdorff, J.M., 2007. The role of higher-level cognitive function in gait: executive dysfunction contributes to fall risk in Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders* 24 (2), 125–137.
- Spreckelmeyer, K.N., Kutas, M., Urbach, T., Altenmuller, E., Munte, T.F., 2009. Neural processing of vocal emotion and identity. *Brain and Cognition* 69 (1), 121–126.
- Stolarova, M., Keil, A., Moratti, S., 2006. Modulation of the C1 visual event-related component by conditioned stimuli: evidence for sensory plasticity in early affective perception. *Cerebral Cortex* 16 (6), 876–887.
- Tallon-Baudry, C., Bertrand, O., Pernier, J., 1999. A ring-shaped distribution of dipoles as a source model of induced gamma-band activity. *Clinical Neurophysiology* 110 (4), 660–665.
- Timm, L., Agrawal, D., Viola, F.C., Sandmann, P., Debener, S., Buchner, A., et al., 2012. Temporal feature perception in cochlear implant users. *PLoS One* 7 (9), e45375.
- Vanderploeg, R.D., Brown, W.S., Marsh, J.T., 1987. Judgments of emotion in words and faces: ERP correlates. *International Journal of Psychophysiology* 5 (3), 193–205.
- Viola, F.C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., Debener, S., 2009. Semi-automatic identification of independent components representing EEG artifact. *Clinical Neurophysiology* 120 (5), 868–877.
- Viola, F.C., Thorne, J.D., Bleeck, S., Eyles, J., Debener, S., 2011. Uncovering auditory evoked potentials from cochlear implant users with independent component analysis. *Psychophysiology* 48 (11), 1470–1480.
- Wagener, K., Brand, T., Kollmeier, B., 1999. Entwicklung und Evaluation eines Satztests in deutscher Sprache III: Evaluation des Oldenburger Satztests. *Zeitschrift für Audiologie* 38 (3), 86–95.
- Wittfoth, M., Schroder, C., Schardt, D.M., Dengler, R., Heinze, H.J., Kotz, S.A., 2010. On emotional conflict: interference resolution of happy and angry prosody reveals valence-specific effects. *Cerebral Cortex* 20 (2), 383–392.
- Yordanova, J., Kolev, V., Demiralp, T., 1997. Effects of task variables on the amplitude and phase-locking of auditory gamma band responses in human. *The International Journal of Neuroscience* 92 (3–4), 241–258.
- Yordanova, J., Kolev, V., Rosso, O.A., Schurmann, M., Sakowitz, O.W., Ozgoren, M., et al., 2002. Wavelet entropy analysis of event-related potentials indicates modality-independent theta dominance. *Journal of Neuroscience Methods* 117 (1), 99–109.