EEG gamma-band activity during audiovisual speech comprehension in different noise environments

Yanfei Lin · Baolin Liu · Zhiwen Liu ·
Xiaorong Gao

Received: 28 August 2014 / Revised: 15 January 2015 / Accepted: 23 January 2015 / Published online: 22 February 2015
© Springer Science+Business Media Dordrecht 2015

Abstract The presence of cross-modal stochastic resonance in different noise environments has been proved in previous behavioral and event-related potential studies, while it was still unclear whether the gamma-band oscillation study was another evidence of cross-modal stochastic resonance. The multisensory gain of gamma-band activity between the audiovisual (AV) and auditory-only conditions in different noise environments was analyzed. Videos of face motion articulating words concordant with different levels of pink noise were used as stimuli. Signal-to-noise ratios (SNRs) of 0, −4, −8, −12 and −16 dB were selected to measure the speech recognition accuracy and EEG activity for 20 healthy subjects. The power and phase of EEG gamma-band oscillations increased in a time window of 50–90 ms. The multisensory gains of evoked and total activity, as well as phase-locking factor, were greatest at the −12 dB SNR, which were consistent with the behavioral result. The multisensory gain of gamma-band activity showed an inverted U-shaped curve as a function of SNR. This finding confirmed the presence of cross-modal stochastic resonance. In addition, there was a significant correlation between evoked activity and phase-locking factor of gamma-band at five different

SNRs. Gamma-band oscillation was believed to play a role in the rapid processing and information linkage strengthening of AV modalities in the early stage of cognitive processes.

Keywords Cross-modal stochastic resonance · Audiovisual speech · Noise · Gamma-band activity · Multisensory integration

Introduction

The cross-modal stochastic resonance (SR) is a nonlinear phenomenon whereby a weak signal response from visual pathway can be enhanced by the noise from the auditory pathway, and the signal response varies as an inverted U-shaped function of the intensity of noise (Manjarrez et al. 2007). Previous behavioral and event-related potential (ERP) studies showed that the SR-like phenomenon involves not only auditory-visual interaction (Ross et al. 2007a, b; Manjarrez et al. 2007; Liu et al. 2013), but also visual-somatosensory (Ramos-Estebanez et al. 2007), auditory-tactile, auditory-proprioceptive, tactile-visual, and tactile-auditory interactions (Lugo et al. 2008a, b) in humans.

Many researchers have investigated auditory-visual speech comprehension in noise environments. Visual articulatory information in noise environments can improve speech recognition accuracy (MacLeod and Summerfield 1987; Sumbly and Pollack 1954). This observation is called multisensory integration. In behavioral studies, Ross et al. (2007a, b) reported that speech recognition performance was measured at −24, −20, −16, −12, −8, −4 and 0 dB signal-to-noise ratios (SNRs) in normal subjects and schizophrenia patients. Results showed that the gain as the

Y. Lin · X. Gao (✉)
School of Medicine, Tsinghua University, Beijing 100084,
People's Republic of China
e-mail: gxr-dea@mail.tsinghua.edu.cn;
gxr-dea@tsinghua.edu.cn

Y. Lin · Z. Liu
School of Information and Electronics, Beijing Institute of
Technology, Beijing 100081, People's Republic of China

B. Liu
School of Computer Science and Technology, Tianjin
University, Tianjin 300072, People's Republic of China

difference in speech recognition accuracies between AV and auditory-only (A) conditions occurred greatest at -12 dB SNR. Electrophysiological studies using electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) have investigated multisensory integration in single-level noise environment (Winneke and Phillips 2011; Callan et al. 2001, 2003; Bishop and Miller 2009). Less electrophysiological research, however, has been conducted on multisensory integration in different noise environments. Liu et al. (2013) tested the audiovisual speech accuracy and recorded EEG signals at five levels of SNRs: -16 , -12 , -8 , -4 and 0 dB. The behavioral result showed that the multisensory gain as the difference in speech recognition accuracies between AV and A conditions was largest at the -12 dB SNR, which was consistent with findings by Ross et al. (2007a, b). Their ERP results revealed that the multisensory gain showed an inverted U-shaped curve as a function of SNR. This observation confirmed the presence of cross-modal SR.

Some studies have indicated that the EEG oscillation power in certain frequency bands predicts perceptual performance (Naue et al. 2011; Senkowski et al. 2008; Buzsáki and Draguhn 2004). An example of representative oscillation is gamma-band activity, whose oscillatory synchronization plays a key role in brain network integration and information selection (Engel et al. 2001; Hermann et al. 2004; Jensen et al. 2007). Previous studies have shown that gamma-band activity was related to attentional selection (Wyart and Tallon-Baudry 2008; Bauer et al. 2006), memory processes (Kaiser et al. 2008; Kaiser and Lutzenberger 2005; Busch et al. 2008), object representation (Tallon-Baudry et al. 1996) and language processes (Bastiaansen and Hagoort 2006). These studies indicated that oscillatory synchronization of gamma-band activity was associated with cognitive abilities. Thus, the investigation of gamma-band oscillation is an important approach to study multisensory integration in different levels of noise environments.

Liu et al. (2013) demonstrated the presence of cross-modal SR by ERP study, while it was still unclear whether the study of gamma-band activity could reveal the cross-modal SR phenomenon. The present study analyzed EEG data in five different noise environments to determine whether or not multisensory gain as the difference in gamma-band activities between AV and A conditions varied as a function of SNR, and whether or not this variation may confirm the presence of cross-modal SR.

Materials and methods

The participants, materials, experimental procedure and EEG-recording were described by Liu et al. (2013) in

detail. The present study used EEG data obtained under AV, A and visual-only (V) conditions to analyze gamma-band activity.

Participants

Twenty volunteers (10 males and 10 females) between the ages of 19 and 29 years old (mean = 24.0, standard deviation = 2.2) participated in this experiment. All participants had normal or corrected-to-normal vision and normal hearing, were without neurological diseases or mental disorders, and were native Chinese speakers and right-handed. The experiment was carried out in accordance with the Declaration of Helsinki and approved by the Tsinghua University Institutional Review Board. All subjects signed the Researchers' Consent Form prior to the experiment and were financially rewarded for their participation.

Materials

The face of a female speaker articulating 250 Chinese words was digitally video-recorded. Each of the 250 video clips was edited using Adobe Premiere Pro CS3 with a length of 667 ms and 20 video frames. All the video clips adopted the NTSC standard with a luminance resolution of 720×480 pixels and a frame rate of 29.97 fps (frame per second). The speech sound of each video clip was extracted, and then the noise reduction and normalization were conducted by Adobe Audition 3.0. These 250 words were separated into five blocks, and five different levels of pink noise were generated by matlab software and added to the corresponding blocks. The intensities of the five different levels of pink noise were determined by the ratio of the variance of each speech signal to the variance of pink noise: 0 , -4 , -8 , -12 and -16 dB, respectively. Each speech word was presented at an intensity of 50 dB SPL measured by the Norsonic 140 Sound level Meter, and the intensities of five different levels of pink noise were presented at 50, 54, 58, 62 and 66 dB SPL.

The experiments comprised three conditions: A, V, and AV, and included 250 stimuli for each condition. A total of 750 stimuli were generated. In the A condition, the speech sounds and different levels of pink noise were synthesized as A stimuli using Matlab software. In the V condition, the silent video of the female speaker articulating the words and different levels of pink noise were synthesized as V video stimuli by Adobe Premiere Pro CS3. In the AV condition, the silent video of the female speaker articulating the words and the corresponding A stimuli were synthesized as audio-visual video stimuli by Adobe Premiere Pro CS3.

Procedure

Participants sat 1 m away from a computer screen at a 6.74° horizontal angle of view and 4.5° vertical angle of view. Sounds came from two loud speakers beside both sides of the screen. Figure 1 shows the experimental procedures under A, AV, and V conditions. The total 750 stimuli were pseudo-randomized. At the beginning of each stimulus, a white cross was first presented randomly for 200–300 ms. If the stimulus was A, a blank screen was shown for 450 ms and then the speech sound mixed pink noise was presented with a black screen. If the stimulus was AV, the first still image of each audio-visual video stimulus was presented for 450 ms and then the speaker's face articulated words concordant with A stimuli was shown. If the stimulus was V, the first still image of each V video stimulus was presented for 450 ms and then the speaker's face articulated words concordant with only the pink noise was shown. The subject was expected to respond as soon as he or she heard the word. Finally, a black screen was displayed for 1,000 ms, and then the next stimulus was presented.

EEG recording

EEG data were recorded using Neuroscan Synamps with 64-channel Ag/AgCl electrodes. The sampling rate was 1,000 Hz, and the right mastoid was online-referenced. The electrode impedance was less than $5\text{ K}\Omega$.

Data analysis

EEG data were preprocessed and analyzed using wavelet transform. Then, the evoked activity, total activity and

phase-locking factor of gamma-band activity were calculated in the time–frequency domain. Finally, multisensory gains as the differences in evoked activity, total activity and phase-locking factor between AV and A conditions were obtained and statistically analyzed (Naue et al. 2011; Başsar-Eroglu et al. 1996). The specific steps are described as follows.

Pre-processing

EEG data were filtered using a band-pass filter from 0.5 to 120 Hz and a 50 Hz notch filter (band width of 4 Hz). The average of the left and right mastoid was used as an offline reference point. The beginning of auditory noise stimuli represented the onset of target stimuli. The EEG data were segmented into single trials spanning from 200 ms before the auditory onset to 400 ms after the auditory onset, with the interval from 200 to 50 ms before the auditory onset calculated as the baseline. EEG epochs were eliminated with a threshold of $\pm 100\text{ }\mu\text{V}$ contaminated by ocular movements or other artifacts. Finally, three participants with low rates of remaining epochs ($<30\%$) were rejected.

Time–frequency analysis

The wavelet transform was computed by convolving the EEG data with the complex Gaussian wavelet. Wavelet analysis was performed for the frequency range of 1–60 Hz with a step of 1 Hz. At 40 Hz, the duration ($2\sigma_t$) and spectral bandwidth ($2\sigma_f$) of the wavelet were 30 ms and 21.2 Hz, respectively. The low gamma-band activity of the EEG data was analyzed at a frequency band of 30–50 Hz.

Analysis of oscillatory activity

In the time–frequency domain, the power and phase of the oscillation activity were analyzed. The power reflected the intensity of the oscillation activity, and the phase indicated the synchronization of oscillation activity. Oscillation activity consisted of evoked and induced responses (Senkowski et al. 2008; Başsar-Eroglu et al. 1996). Many researchers have performed investigations of oscillation activity (Naue et al. 2011; Tallon-Baudry et al. 1996, 1998; Yuval-Greenberg and Deouell 2007; Fründ et al. 2007a; Demiralp et al. 2007). As previously described, the evoked and total activities as well as the phase-locking factor were traditionally calculated to analyze the power and phase of gamma-band activity. We assumed that the trial number of EEG data was defined by n , that the time response of each EEG trial was $x(t)$, and that the function of wavelet transform was $CWT(x)$. Then the evoked activity $E(t, f)$, total activity $T(t, f)$ and phase-locking factor $P(t, f)$ are expressed as follows.

Fig. 1 Experimental procedures

1. The evoked activity was used to analyze the evoked response whose phase locked to the stimulus. The activity was obtained by wavelet transform of the averaged ERP and calculation of the absolute value of the result (Naue et al. 2011; Tallon-Baudry et al. 1996; Fründ et al. 2007a).

$$E(t, f) = \left| CWT \left(\frac{1}{n} \sum_{i=1}^n x_i(t) \right) \right| \quad (1)$$

2. The total activity included the phase-locked and non-phase-locked responses. The activity was obtained by wavelet transform of each trial and calculation of the absolute value. Finally, the transformation result was averaged (Naue et al. 2011; Tallon-Baudry et al. 1996; Fründ et al. 2007a).

$$T(t, f) = \frac{1}{n} \sum_{i=1}^n |CWT(x_i(t))| \quad (2)$$

3. Phase-locking factor was used to analyze the phase without regard to the amplitude. Wavelet transform was first applied across trials, and the amplitude of the time–frequency representation was normalized and averaged. Finally, the absolute value of the result was calculated. The phase-locking factor ranged from 0 to 1, where 0 indicates the phases across trials cancel each other out, and 1 represents phase-locking across trials (Naue et al. 2011; Tallon-Baudry et al. 1996; Fründ et al. 2007a).

$$P(t, f) = \left| \frac{1}{n} \sum_{i=1}^n \frac{CWT(x_i(t))}{|CWT(x_i(t))|} \right| \quad (3)$$

4. The analyzed time-window was defined by the time-period during which the power and phase of gamma-band activity increased. The resulting time–frequency representation in Fig. 2 shows that the enhancements of gamma-band power and phase were mainly concentrated between 50 and 90 ms after auditory onset (cf. the red boxes in Fig. 2). A region of interest exhibited the strongest response to auditory stimulation, and nine corresponding electrode sites were identified: F1, Fz, F2, FC1, FCz, FC2, C1, Cz and C2. Then, multisensory gains as the differences in the evoked and total activities as well as the phase-locking factor between AV and A conditions were calculated at five levels of SNRs.
5. Repeated-measures ANOVA was performed to determine whether or not SNRs exert effects on multisensory gains as the differences in the evoked and total activities as well as the phase-locking factor between AV and A conditions. Pair-wise follow-up analysis with a least significant difference test was performed to

determine whether or not a significant difference in the multisensory gain may be observed among the different levels of SNRs. Geisser-Greenhouse correction was used to evaluate statistical effects with more than one degree of freedom in the numerator. Fitted curves of the multisensory gains using sine function were conducted to analyze the maximum gain of multisensory integration. Significant difference (p value) was set to 0.05.

Results

Gamma-band activity results

The time responses of the evoked and total activities as well as the phase-locking factor for the FCz electrode at -12 dB SNR are shown in Fig. 2. Peaks appeared in the time window of approximately 50–90 ms. Topographic maps of the evoked activities and the phase-locking factor in the time-window of 50–90 ms at -12 dB SNR are plotted in Fig. 3a, b. The total activity was weak and not shown. The evoked activity and phase-locking factor were distributed in the similar brain regions. Under the AV condition those were mainly concentrated in the frontal, fronto-central and central regions. Under the A condition those were mainly distributed in the left hemisphere, such as fronto-central, central and parietal regions. Under the V condition those were mainly concentrated in frontal, fronto-central and central regions. According to the distributions of brain regions for evoked activity and phase-locking factor at -12 dB SNR, three electrodes were selected to analyze the correlation between evoked activity and phase-locking factor at five different SNRs. They were electrode Fz under AV condition, electrode C5 under A condition, and electrode F2 under V condition. Correlation coefficient between evoked activity and phase-locking factor at -12 dB SNR was conducted. The results in Fig. 3c show that evoked activity of gamma-band are significantly correlated with phase-locking factor (Corrs > 0.750 , $ps < 0.001$) at -12 dB SNR. Next, evoked activity and phase-locking factor at five SNRs were analyzed, as shown in Fig. 3d, e. Correlation coefficient between evoked activity and phase-locking factor at five SNRs was calculated. The results in Fig. 3e show that there is a significant correlation between evoked activity and phase-locking factor of gamma-band activities for five SNRs (Corrs ≥ 0.931 , $ps \leq 0.022$).

Multisensory gains as the differences in evoked and total activities as well as phase-locking factor between AV and A conditions at five levels of SNRs in the time window of

Fig. 2 Evoked activity, total activity and phase-locking factor of the FCz electrode at -12 dB SNR under AV (a–c), A (d–f) and V (g–i) conditions

50–90 ms are illustrated in Fig. 4. The multisensory gain was largest at -12 dB SNR.

For the evoked activity, repeated-measures ANOVA of SNRs (five levels) $\times$ electrodes (nine levels: F1, Fz, F2, FC1, FCz, FC2, C1, Cz and C2) showed a significant effect of SNRs on the multisensory gains [$F(4,576) = 9.704$, $p = 0.000 < 0.001$]. Pair-wise follow-up analysis revealed that the multisensory gain of evoked activity at -12 dB SNR was significantly higher than those at 0, -4 , -8 and -16 dB SNRs ($ps < 0.006$).

For the total activity, repeated-measures ANOVA of SNRs (five levels) $\times$ electrodes (nine levels: F1, Fz, F2, FC1, FCz, FC2, C1, Cz and C2) showed a significant effect of SNRs on the multisensory gains [$F(4,576) = 12.514$, $p = 0.000 < 0.001$]. Pair-wise follow-up analysis revealed that the multisensory gain at -12 dB SNR was significantly higher than those at 0, -4 , -8 and -16 dB SNRs ($p = 0.000 < 0.001$).

For the phase-locking factor, repeated-measures ANOVA of SNRs (five levels) $\times$ electrodes (nine levels: F1, Fz, F2, FC1, FCz, FC2, C1, Cz and C2) showed a significant effect of SNRs on the multisensory gains [$F(4,576) = 11.306$, $p = 0.000 < 0.001$]. Pair-wise follow-up analysis revealed that the multisensory gain of phase-locking factor at -12 dB SNR was significantly

higher than those at -4 , -8 and -16 dB SNRs ($ps = 0.000 < 0.01$), and that no significant difference in multisensory gain of phase-locking factor was found between -12 and 0 dB SNRs ($p = 0.063 > 0.05$). Fitted curve for the multisensory gain of phase-locking factor was performed using the sine function, as shown in Fig. 4d. The result indicates that the red line represents a best-fit regression with a maximum R-square value of 0.96, and that the multisensory gain at -12 dB SNR is much higher than the predicted value comparing that at 0 dB SNR. Therefore, the multisensory gain of phase-locking factor was greatest at -12 dB SNR.

Discussion

Cross-modal stochastic resonance

Gamma-band activities showed that multisensory gains as the differences in evoked and total activities as well as phase-locking factor between AV and A conditions were largest at -12 dB SNR, similar to the behavioral results (Liu et al. 2013). Therefore, the present results further proved that gamma-band oscillation could predict cognitive functions, which was consistent with findings in previous studies

Fig. 3 Topographies of evoked activity (a) and phase-locking factor (b) are plotted at -12 dB SNR under AV, A and V conditions. Correlation coefficient between evoked activity and phase-locking factor at -12 dB SNR are conducted for Electrode Fz under AV condition, for Electrode C5 under A condition, and for Electrode F2 under V condition, as shown in (c). Each circle in subgraph

corresponds to evoked activity and phase-locking factor of each subject. Evoked activity and phase-locking factor at five different SNRs are shown in (d) and (e). The error bars represent the standard error of the mean (SEM). Correlation coefficient between evoked activity and phase-locking factor at five SNRs are calculated in (e). *Corr* is correlation coefficient and *p* is statistically significant value

Fig. 4 Gains (AV-A) of evoked activity (a), total activity (b) and phase-locking factor (c) at 0, -4, -8, -12 and -16 dB SNRs. The error bars represent the standard error of the mean (SEM). Fitted curves of the multisensory gain of phase-locking factor in (d) obtained using the sine function. The blue line represents the fitted curve for all five data points. The red line shows the fitted curve for four data points without the data point at -12 dB SNR. The magenta line illustrates the fitted curve for four data points without the data point at 0 dB SNR

(Spencer et al. 2003, 2004). Furthermore, the multisensory gains of evoked and total activities as well as phase-locking factor showed an inverted U-shaped curve as a function of SNR, which agreed well with the ERP results of Liu et al. (2013). Thus, the observations of gamma-band oscillation confirmed the presence of cross-modal SR.

Latency of gamma-band activity

The evoked activity, total activity and phase-locking factor of gamma-band activity under AV and A conditions increased in the time window of 50–90 ms. These results were similar to those in previous studies. The latency of gamma-band activity appeared in the time window of 30–120 ms in Senkowski's study (Senkowski et al. 2007). Early gamma-band responses of evoked, total activities and phase-locking factor occurred at the latency of 60–120 ms in Fründ's study (Fründ et al. 2007b).

Comparison with ERP gain

In the present study, the multisensory gains of gamma-band evoked and total activities as well as phase-locking factor appeared in the time window of 50–90 ms, while the multisensory ERP gain was in the time window of 130–200 ms as shown in Liu et al. (2013). Gamma-band activity during multisensory integration was much earlier than ERP. Early stages of ongoing gamma-band activity can be modulated by top-down processes (Ohl and Scheich 2005). Early gamma-band activity may adjust the neural network of different sensory cortices to respond synchronously (Wespatat et al. 2004; Paik et al. 2009).

Correlation between evoked activity and phase-locking factor

There was a significant correlation between evoked activity and phase-locking factor of gamma-band at five different

SNRs. Evoked gamma-band activity is phase-locked to the onset of the stimulus (Herrmann et al. 2010). Phase-locking factor indicates the degree of the inter-trial signal coherence. A higher phase-locking factor represents perfect phase locking across trials (Naue et al. 2011; Schneider et al. 2008). Both evoked activity and phase-locking factor are obtained by averaging all trials. The former includes power and phase, and the latter normalizes amplitude and only considers phase. Both evoked activity and phase-locking factor represent phase-locking to stimulus onset. The present study analyzed more conditions (5 SNRs), therefore the significant correlation between evoked activity and phase-locking factor is more evident. However, other studies that analyzed one or two conditions did not report similar results (Stevenson et al. 2012; Tallon-Baudry et al. 1996; Fründ et al. 2007a).

Auditory noise

In the present paper, pink noise but not white noise was employed based on two considerations. First, pink noise could be used to simulate the background noise of brain (Lin et al. 2013). The second, we just referred to the experiment method of audio-visual speech in Ross et al. 2007a, b) to adopt the pink noise. Some studies about SR indicated that some specific noise-bandwidths [e.g. in multisensory sensorimotor integration (Trenado et al. 2014), or in sensory neurons (Nozaki et al. 1999)] were more effective to produce a more powerful SR. In the future work, we will try to use these specific noise-bandwidths in audio-visual speech comprehension to determine whether or not EEG gamma-band varies as a function of SNR.

Other oscillation activities

In the present study, other oscillation activities such as theta-band (4–8, 280–460 ms), alpha-band (8–13, 260–400 ms), and beta-band (13–30, 260–320 ms) were also analyzed (cf. the white boxes in Fig. 2). We did not obtain significantly higher oscillation activities at -12 dB SNR than at the other four SNRs, nor did we find any inverted U-shaped curves as a function of SNR. However, the multisensory gain of gamma-band oscillation showed an inverted U-shaped curve as a function of SNR, which was consistent with the behavioral results. It was indicated that higher cognitive processing was correlated with the gamma-band oscillation. In addition, for other oscillation activities, no inverted U-shaped curves as a function of SNR were found. We did not know whether the reason was related with the use of pink noise. In the future study, we will exploit the white noise instead of pink noise to study whether the other EEG bands could respond differently.

Related explanation

Earlier phase-locking activities, such as evoked activity and phase-locking factor, were associated with modulations of memory matches (Herrmann et al. 2010, 2004; Kayser et al. 2008; Tallon-Baudry et al. 1998) and attention or arousal (Wyart and Tallon-Baudry 2008; Bauer et al. 2006). The phase of gamma-band oscillation could be spontaneously reset to respond to new stimuli. Each stimulus triggers a spike in neuron activity that could be associated with gamma-band phase-locking (Fries et al. 2007). Phase-resetting in different cortices could simultaneously occur and quickly reset to a given value determined by the stimulus properties (Kazantsev et al. 2004). Phase-locking to the stimuli would allow neural cells to achieve phase synchronization. Therefore, phase-resetting of gamma-band oscillation plays a role in the rapid processing and linkage strengthening of different sensory information.

Clinical applications

Now, more studies pay more attention to Autism Spectrum Disorders (ASD) (Matsuura et al. 2008; Mirman et al. 2012) and Attention Deficit Hyperactivity Disorder (ADHD) (Duch and Dobosz 2011) patients. Previous studies showed that ASD persons had weak functional connections between the frontal lobe and the rest of the cortex (Murias et al. 2007), and a different brain structure in the prefrontal, striatal, cerebellar and thalamic circuits was found in ADHD individuals compared to controls (Krain and Castellanos 2006). Gamma-band oscillatory synchronization plays a key role in brain network integration and attention selection (Engel et al. 2001; Jensen et al. 2007). Therefore, EEG gamma-band analysis during audio-visual speech comprehension may be a promising diagnostic tool to assess the functional status of brain in such as ADHD and ASD.

Conclusion

This study investigated the gamma-band activity of multisensory integration in different noise environments. Speech recognition accuracy and EEG signals were measured at 0, -4 , -8 , -12 and -16 dB SNRs. The evoked and total activities as well as the phase-locking factor of gamma-band activity were analyzed. Gamma-band activity was enhanced in the time window of 50–90 ms. The multisensory gains of the evoked and total activities as well as the phase-locking factor were greatest at -12 dB SNR, which was consistent with the behavioral results. The multisensory gain of gamma-band oscillation showed an inverted U-shaped curve as a function of SNR. This results

confirmed the presence of cross-modal SR. In addition, there was a significant correlation between evoked activity and phase-locking factor of gamma-band at five different SNRs. Gamma-band activity played a role in rapid processing and linkage strengthening of audiovisual modalities in early cognition.

Acknowledgments This work was supported by National Basic Research Program (973) of China (Nos. 2011CB933204, 2013CB329301), National Natural Science Foundation of China (Nos. 61271128, 91120007, 61271112), China Postdoctoral Science Foundation funded project and Chinese 863 Project (No. 2012AA011601).

References

- Basşar-Eroglu C, Strüber D, Schürmann M, Stadler M, Başşar E (1996) Gamma band responses in the brain: a short review of psychophysiological correlates and functional significance. *Int J Psychophysiol* 24:101–112
- Bastiaansen M, Hagoort P (2006) Oscillatory neuronal dynamics during language comprehension. *Prog Brain Res* 159:179–196
- Bauer M, Oostenveld R, Peeters M, Fries P (2006) Tactile spatial attention enhances gamma-band activity in somatosensory cortex and reduces low-frequency activity in parieto-occipital areas. *J Neurosci* 26:490–501
- Bishop CW, Miller LM (2009) A multisensory cortical network for understanding speech in noise. *J Cogn Neurosci* 21:1790–1805
- Busch NA, Groh-Bordin C, Zimmer HD, Herrmann CS (2008) Modes of memory: early electrophysiological markers of repetition suppression and recognition enhancement predict behavioral performance. *Psychophysiology* 45:25–35
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304:1926–1929
- Callan DE, Callan AM, Kroos C, Vatikiotis-Bateson E (2001) Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. *Cogn Brain Res* 10:349–353
- Callan DE, Jones JA, Munhall K, Callan AM, Kroos C, Vatikiotis-Bateson E (2003) Neural processes underlying perceptual enhancement by visual speech gestures. *Neuro Report* 14:2213–2218
- Demiralp T, Bayraktaroglu Z, Lenz D, Junge S, Busch N, Maess B, Ergen M, Herrmann C (2007) Gamma amplitudes are coupled to theta phase in human EEG during visual perception. *Int J Psychophysiol* 64:24–30
- Duch W, Dobosz K (2011) Visualization for understanding of neurodynamical systems. *Cogn Neurodyn* 5:145–160
- Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* 2:704–716
- Fries P, Nikolić D, Singer W (2007) The gamma cycle. *Trends Neurosci* 30:309–316
- Fründ I, Schadow J, Busch NA, Körner U, Herrmann CS (2007a) Evoked gamma oscillations in human scalp EEG are test-retest reliable. *Clin Neurophysiol* 118:221–227
- Fründ I, Busch NA, Körner U, Schadow J, Herrmann CS (2007b) EEG oscillations in the gamma and alpha range respond differently to spatial frequency. *Vision Res* 47:2086–2098
- Herrmann CS, Munk MH, Engel AK (2004) Cognitive functions of gamma-band activity: memory match and utilization. *Trends Cogn Sci* 8:347–355
- Herrmann CS, Fründ I, Lenz D (2010) Human gamma-band activity: a review on cognitive and behavioral correlates and network models. *Neurosci Biobehav Rev* 34:981–992
- Jensen O, Kaiser J, Lachaux JP (2007) Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci* 30:317–324
- Kaiser J, Lutzenberger W (2005) Cortical oscillatory activity and the dynamics of auditory memory processing. *Rev Neurosci* 16:239–254
- Kaiser J, Heidegger T, Wibral M, Altmann CF, Lutzenberger W (2008) Distinct gamma-band components reflect the short-term memory maintenance of different sound lateralization angles. *Cereb Cortex* 18:2286–2295
- Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. *Cereb Cortex* 18:1560–1574
- Kazantsev VB, Nekorkin VI, Makarenko VI, Llinás R (2004) Self-referential phase reset based on inferior olive oscillator dynamics. *Proc Natl Acad Sci USA* 101:18183–18188
- Krain AL, Castellanos FX (2006) Brain development and ADHD. *Clin Psychol Rev* 26:433–444
- Lin Y, Wu W, Wu C, Liu B, Gao X (2013) Extraction of mismatch negativity using a resampling-based spatial filtering method. *J Neural Eng* 10:026015
- Liu BL, Lin YF, Gao XR, Dang JW (2013) Correlation between audio-visual enhancement of speech in different noise environments and SNR: a combined behavioral and electrophysiological study. *Neuroscience* 247:145–151
- Lugo JE, Doti R, Faubert J (2008a) Ubiquitous crossmodal stochastic resonance in humans: auditory noise facilitates tactile, visual and proprioceptive sensations. *PLoS ONE* 3:e2860
- Lugo JE, Doti R, Wittich W, Faubert J (2008b) Multisensory integration central processing modifies peripheral systems. *Psychol Sci* 19:989–997
- MacLeod A, Summerfield Q (1987) Quantifying the contribution of vision to speech perception in noise. *Br J Audiol* 21:131–141
- Manjarrez E, Mendez I, Martinez L, Flores A, Mirasso CR (2007) Effects of auditory noise on the psychophysical detection of visual signals: cross-modal stochastic resonance. *Neurosci Lett* 415:231–236
- Matsuura H, Tateno K, Aou SJ (2008) Dynamical properties of the two-process model for sleep-wake cycles in infantile autism. *Cogn Neurodyn* 2:221–228
- Mirman D, Irwin JR, Stephen DG (2012) Eye movement dynamics and cognitive self-organization in typical and atypical development. *Cogn Neurodyn* 6:61–73
- Murias M, Webb SJ, Greenson J, Dawson G (2007) Resting state cortical connectivity reflected in EEG coherence in individuals with autism. *Biol Psychiatry* 62:270–273
- Naue N, Rach S, Strüber D, Huster RJ, Zaehle T, Körner U, Herrmann CS (2011) Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *J Neurosci* 31:7729–7736
- Nozaki D, Mar D, Grigg P, Collins J (1999) Effects of colored noise on stochastic resonance in sensory neurons. *Phys Rev Lett* 82:2402–2405
- Ohl FW, Scheich H (2005) Learning-induced plasticity in animal and human auditory cortex. *Curr Opin Neurobiol* 15:470–477
- Paik S, Kumar T, Glaser D (2009) Spontaneous local gamma oscillation selectively enhances neural network responsiveness. *PLoS Comput Biol* 5:e1000342
- Ramos-Estebanez C, Merabet LB, Machii K, Fregni F, Thut G, Wagner TA, Romei V, Amedi A, Pascual-Leone A (2007) Visual phosphene perception modulated by subthreshold cross-modal sensory stimulation. *J Neurosci* 27:4178–4181
- Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007a) Do you see what I am saying? Exploring visual enhancement of

- speech comprehension in noisy environments. *Cereb Cortex* 17:1147–1153
- Ross LA, Saint-Amour D, Leavitt VM, Molholm S, Javitt DC, Foxe JJ (2007b) Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophr Res* 97:173–183
- Schneider TR, Debener S, Oostenveld R, Engel AK (2008) Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *NeuroImage* 42:1244–1254
- Senkowski D, Talsma D, Grigutsch M, Herrmann CS, Woldorff MG (2007) Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia* 45:561–571
- Senkowski D, Schneider TR, Foxe JJ, Engel AK (2008) Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci* 31:401–409
- Spencer KM, Nestor PG, Niznikiewicz MA, Salisbury DF, Shenton ME, McCarley RW (2003) Abnormal neural synchrony in schizophrenia. *J Neurosci* 23:7407–7411
- Spencer KM, Nestor PG, Perlmuter R, Niznikiewicz MA, Klump MC, Frumin M, Shenton ME, McCarley RW (2004) Neural synchrony indexes disordered perception and cognition in schizophrenia. *Proc Natl Acad Sci USA* 101:17288–17293
- Stevenson RA, Bushmakina M, Kim S, Wallace MT, Puce A, James TW (2012) Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topogr* 25:308–326
- Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. *J Acoust Soc Am* 26:212–215
- Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci* 16:4240–4249
- Tallon-Baudry C, Bertrand O, Peronnet F, Pernier J (1998) Induced gamma-band activity during the delay of a visual short-term memory task in humans. *J Neurosci* 18:4244–4254
- Trenado C, Mikulic A, Manjarrez E, Mendez-Balbuena I, Schulte-Monting J, Huethe F, Hepp-Reymond MC, Kristeva R (2014) Broad-band Gaussian noise is most effective in improving motor performance and is most pleasant. *Front Hum Neurosci* 8:22
- Wespatat V, Tennigkeit F, Singer W (2004) Phase sensitivity of synaptic modifications in oscillating cells of rat visual cortex. *J Neurosci* 24:9067–9075
- Winneke AH, Phillips NA (2011) Does audiovisual speech offer a fountain of youth for old ears? An event-related brain potential study of age differences in audiovisual speech perception. *Psychol Aging* 26:427–438
- Wyart V, Tallon-Baudry C (2008) Neural dissociation between visual awareness and spatial attention. *J Neurosci* 28:2667–2679
- Yuval-Greenberg S, Deouell Y (2007) What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. *J Neurosci* 27:1090–1096