Bio-Medical Materials and Engineering 26 (2015) S767–S778
DOI 10.3233/BME-151368
IOS Press

Temporal integration reflected by frequency following response in auditory brainstem[1]

Qin Xu[a,b] and Datian Ye[a,b,*]
[a]*Department of Biomedical Engineering, Tsinghua University, Beijing, 100084, China*
[b]*Research Center of Biomedical Engineering, Graduate School at Shenzhen Campus of Tsinghua University, Shenzhen 518055, China*

Abstract. Auditory temporal integration (ATI) has been widely described in psychoacoustic studies, especially for loudness perception. Loudness increases with increasing sound duration for durations up to a time constant about 100 ~ 200 ms, and then loudness becomes saturated with more duration increase. However, the electrophysiological mechanism underlying the ATI phenomenon has not been well understood. To investigate ATI at the brainstem level of auditory system and its relationship to cortical and behavioral ATI, frequency follow response (FFR) was acquired in our study. Simultaneously, ATI in auditory cortex was evaluated by cortical response P1. Behavioral loudness and electrophysiological measures were estimated from normal-hearing young adults for vowel /a/ whose durations varied from 50 ms to 175 ms. Significant effects of stimulus duration were found both on FFR and P1 amplitudes. Linear regression analysis revealed that as stimulus duration increased, brainstem FFR amplitude was significantly associated with cortical P1 amplitude and behavioral loudness, which confirmed the existence of temporal integration in auditory brainstem. Moreover, behavioral loudness ATI was better predicted using brainstem and cortical measures together than merely using each one separately, indicating an interplay and coordination for ATI across the three levels along auditory pathway.

Keywords: Auditory temporal integration (ATI), frequency following response (FFR), cortical response P1, loudness

1. Introduction

An important parameter influencing the perception of a sound is duration. Psychoacoustic studies has shown that behavioral threshold decreases and loudness increases as stimulus duration grows from a few milliseconds to a time constant (100 ~ 200 ms) [1, 2]. Behavioral threshold or loudness becomes independent of stimulus duration beyond the time constant. The improvement in threshold and increment in loudness along with increasing duration is attributed to a process of temporal integration in auditory system. Auditory temporal integration (ATI) is a fundamental ability of auditory system, which should also be reflected electrophysiologically.

However, the electrophysiological mechanism of ATI is a matter of controversy, concerning uncertainty about the sites of ATI. Electrophysiological correlates of ATI were addressed in both magnetoencephalography (MEG) and electroencephalography (EEG) studies. Effects of stimulus

[1] The authors report no declarations of interest upon this paper.

[*] Address for correspondence: Datian Ye, Research Center of Biomedical Engineering, Graduate School at Shenzhen Campus of Tsinghua University, Shenzhen 518055, China. Tel.: 13855197357; E-mail: helloxuqin2@163.com.

duration on auditory evoked potentials (AEP) and auditory evoked fields (AEF) were investigated. Studies on late responses such as P1-N1-P2 waves indicated the existence of ATI at the cortex level, since the response amplitude (especially N1 amplitude) increased with the increment in stimulus duration [3-8]. The amplitudes of late negatives N2 and N4 were also increased for stimuli with longer consonant-vowel transition duration [8]. In addition, the detection threshold for N1-P2 complex decreased with increasing duration [9]. Cortical ATI was also supported by steady-state magnetic field evoked by 40 Hz amplitude modulation tone-bursts, since the response amplitude progressively increased over a time period of 200 ms [7].

Some studies suggested that there was little temporal integration prior to primary auditory cortex [10, 11]. The detection thresholds of auditory brainstem response (ABR) evoked by 2000 Hz tone bursts were not affected by stimulus duration [12]. When evoked by a series of tone pulses at behavioral threshold, ABR to the first pulse was not different from ABR to the last pulse. Similarly, no difference was reported between the middle latency responses (Pa) to the first and the last pulses. Since the origin of wave Pa was primary auditory cortex, it was concluded that ATI originated more centrally than primary auditory cortex [10]. On the contrary, studies on auditory nerve fibers (ANFs) of cats indicated that ATI might take place in the first synapse of the auditory system, between the inner hair cell and the afferent auditory-nerve fiber [13, 14]. The ATI of ANF was remarkably similar to that obtained at the perceptual level [14].

To solve the controversy, it was essential to do more investigation on electrophysiological correlates of ATI. As reviewed above, most AEPs have been used as electrophysiological correlates of ATI, except FFR. FFR is a sustained response recorded from human scalp to reflect the summation of phase-locking activities among neurons at auditory brainstem. Different from other kinds of AEPs, FFR is quite unique by mimicking the stimulus itself with high fidelity [15, 16]. Thus, FFR was expected to be more suitable for evaluating ATI in brainstem compared to ABR which was primarily a transient onset response evoked by clicks or tone-bursts. In previous studies, electrophysiological correlates of ATI at different auditory levels were studied separately and were only compared to behavioral ATIs. It is also necessary and meaningful to examine the relationship between electrophysiological ATIs at different levels of auditory nervous system.

Therefore, in our study we aimed to examine whether ATI existed at the level of brainstem by measuring FFR. If brainstem ATI was confirmed, the relationship between brainstem ATI and cortical ATI would be assessed, as well as the relationships between the two electrophysiological ATIs and behavioral ATI. Specifically, brainstem FFR and cortical responses were acquired simultaneously, evoked by stimuli with durations ranging from 50 ms to 175 ms. Behavioral loudness perceptions to the same stimuli were also estimated. ATI was evaluated through assessing the effects of stimulus duration on the neural and behavioral measures. The relationship between brainstem and cortical ATIs was analyzed, as well as the electrophysiological and behavioral relationships, so as to examine possible interplay and coordination across different auditory levels.

2. Methods

2.1. Subjects

Ten college students of Tsinghua University (4 females, 6 males; aged from 21 to 26 years) participated in this study. All were native Chinese speakers, who exhibited normal-hearing sensitivity (<20 dB HL for octave frequencies of 250 ~ 8000 Hz) and reported no history of neurological or

psychiatric illness. They all gave their informed consent to participate in the study, in compliance with a protocol approved by the institutional review board at Tsinghua University.

2.2. Stimuli

The stimulus was a synthetic vowel /a/ consisted of fifteen harmonic components added in cosine phase. The F0 was 110 Hz, the first formant (F1) was 770 Hz, and the second formant (F2) is 1320 Hz (Figure 1). Six durations were used, which were 50 ms, 75 ms, 100 ms, 125 ms, 150 ms and 175 ms. All stimuli had 5-ms onset/offset ramps with a cos2 window, and had equal overall intensities with a fixed RMS level of 75dB SPL. Stimuli were presented binaurally through insert earphones (Etymotics ER-2, Elk Grove Village, IL). Calibration was performed with a Brüel & Kjær type 3160-A-042 sound analyzer and a type 4157 2cc-coupler. The stimulus *.WAV file was generated with a resolution of 16-bit and a sampling rate of 48 kHz using MATLAB (Mathworks, Natick, MA).

2.3. Behavioral procedures

The cross-modality matching (CMM) procedure was used to obtain loudness function of duration. For CMM procedure, line length was most commonly used to match the loudness of a sound [17-19].

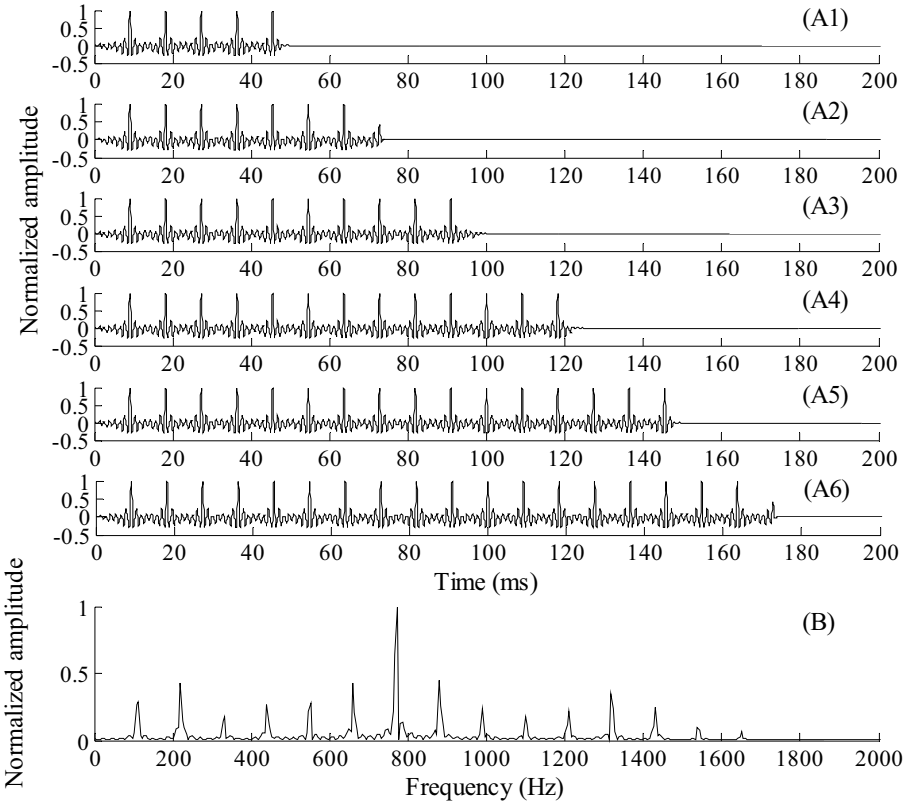

Fig. 1. Stimulus waveforms and spectra. The top six panels (A1-A6) show stimulus waveforms of durations 50 ms, 75 ms, 100 ms, 125 ms, 150 ms and 175 ms. The stimuli have the same spectral distribution, as in the bottom panel (B). The y-axes are normalized amplitudes in the stimulus illustrations, since the actual sound pressure level is determined by the calibration using the whole simulation system.

Each subject was asked to match a line length drawn on a computer screen to the loudness of a sound. Subjects were instructed to match a longer line length to a louder sound.

One block of trials contained six matches for the six stimuli presented randomly. Within a block, the next trial began 2 s after the subject matched a line length. One block of trials was given first as training. Then, ten blocks of trials were completed, that is, ten matches were made for each stimulus. The result of each block was first normalized. Then the final CMM was obtained as the average of the ten values that were matched to each stimulus after normalization.

2.4. Electrophysiological measurements

2.4.1. FFR recording

Six stimuli with different durations were presented in a random order, using the stimulus presentation software Neuroscan Stim 2 (Compumedics; Charlotte, NC, USA). Random order was taken to avoid effects of adaption and nonequivalent background noise. Inter-stimulus interval was in the range of 200~250 ms randomly. For each duration, stimuli of alternating polarities were presented with 1024 tokens per polarity. For one subject, a total of 1024*2*6 tokens were presented. To make sure there was no stimulus artifact in FFR recordings, insert earphones with 580-mm tube were shielded using copper mesh. The effectiveness of shielding was tested by recordings with the earphones tube clamped and also by the latency of FFR responses [20]. The latencies of all subjects' FFRs were around 5~8 ms.

Subjects were seated comfortably in the booth watching silent movies during electrophysiological data acquisition. Responses were recorded using Neuroscan Curry 7.0 (Compumedics; Charlotte, NC, USA) with a sampling rate of 20 kHz and a band-pass filter (0.05 Hz~8 kHz). A vertical electrode montage was used, with non-inverting electrode placed on vertex (+, Cz), reference electrode placed on the right mastoid (-, M2) and the common ground electrode placed on forehead. All electrode impedances were below 3 kΩ. Data processing was conducted offline using MATLAB.

2.4.2. FFR data analysis

Continuous recordings were segmented into sweeps of 290 ms including 40 ms pre-stimulus activity (-40 ms ~ 250 ms) and digitally band-pass filtered using FIR filter (80-3000Hz, 800th order, zero-phase). Sweeps contained voltages greater than $\pm 30\mu V$ were rejected before averaging. Responses to alternating polarities were added to get FFR corresponding to the envelope information [21, 22]. For one subject, about 2000 sweeps were averaged for each duration.

Firstly, temporal measures of FFR latency and amplitude were extracted. FFR peaks were identified visually by the two authors independently, and FFR latency was calculated as the time difference between FFR peaks and stimulus peaks. The portion from 10 ms to the stimulus offset was regarded as the sustained response. The root mean square (RMS) amplitude of the sustained response was calculated, as well as the RMS of residual background noise estimated by the 40-ms pre-stimulus activity. The RMS amplitude of FFR (FFRrms) was corrected by subtracting the noise from the sustained response as in the following equation:

$$FFRrms = \sqrt{Signal_{rms}{}^2 - Noise_{rms}{}^2} \tag{1}$$

where Signalrms was RMS amplitude of the sustained response and Noiserms was RMS of the residual noise. Thus, if Signalrms was smaller than Noiserms, i.e. signal-to-noise ratio smaller than 0

dB, it would be taken as that no FFR existed.

Secondly, amplitudes of individual components at harmonic frequencies in frequency domain were also calculated. Amplitude spectra of the sustained response and the residual noise were obtained using Fast Fourier Transform (FFT) in length of 4096 points with Hanning windowing. Then, the amplitude at a specific harmonic frequency such as nth ($FreqAmp\_H_n$) was calculated as follows:

$$FreqAmp\_H_n = \sqrt{signalFFT^2(H_n) - noiseFFT^2(H_n)} \tag{2}$$

where $signalFFT$ and $noiseFFT$ were spectral amplitudes at nth harmonic (H_n) for the sustained response and residual noise, respectively. No response component was considered to be existed in case of $signalFFT$ smaller than $noiseFFT$ at a given harmonic frequency. Averaged amplitude of the first five harmonics ($H_1 \sim H_5$) was taken as the spectral amplitude for lower frequencies (FFRlowFreqAmp). The amplitudes of the following four harmonics ($H_6 \sim H_9$) were also averaged to represent the spectral amplitude for higher frequencies (FFRhighFreqAmp).

2.4.3. Cortical response P1 analysis

Cortical responses were extracted from the same recordings as brainstem FFR. Continuous recordings were segmented (-200 ~ 350 ms), band-pass filtered (1~30 Hz), baselined to the pre-stimulus interval, and subsequently averaged to obtained cortical responses for each stimulus duration. Sweeps exceeding $\pm 50\mu V$ were rejected before averaging. Peak amplitude and latency were measured for the prominent positivity between 40 and 100 ms, i.e. P1.

2.5. Statistical analysis

In order to determine how the behavioral loudness and neural measures (amplitudes and latencies of FFR and P1) changed with increases in stimulus duration, separate two-way mixed model ANOVAs (with subjects as a random factor and durations as a fixed factor) were used, followed by Bonferroni-corrected post-hoc multiple comparisons. To evaluate the trends of neural response measures with increasing stimulus duration, linear regression analyses were conducted. Grand average of neural measures (including FFR amplitude and latency, P1 amplitude and latency) across all subjects was regarded as the dependent variable, and stimulus duration was taken as the independent variable, respectively. Pairwise linear regressions were performed to assess the relationships between neural measures (FFRrms amplitude and P1 amplitude) and behavioral loudness estimation (normalized loudness by CMM). Grand average of loudness estimation was regarded as the dependent variable, and neural measures were taken as the independent variables, respectively. All statistical analyses were conducted with IBM SPSS Statistics 20.0 software. Shapiro-Wilk tests were applied to ensure that all variables were normally distributed, and Levene's tests were used to ensure homogeneity of variance for all variables.

3. Results

3.1. Brainstem FFR temporal waveforms and spectra

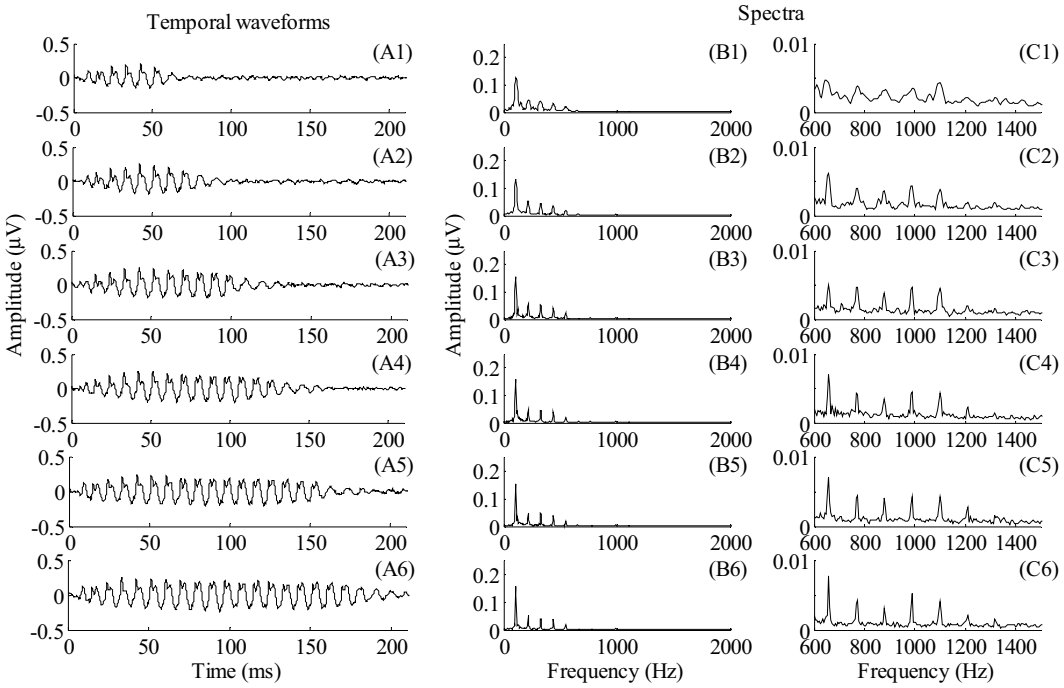

Fig. 2. FFR temporal waveforms and amplitude spectra. Panels in the left column (A1-A6) show the waveforms of FFRs evoked by stimuli with different durations. From top to bottom, the durations are 50 ms, 75 ms, 100 ms, 125 ms, 150 ms and 175 ms, respectively. The corresponding amplitude spectra are illustrated in the middle column (B1-B6), and the enlarged spectral distributions at higher frequencies 600 Hz ~ 1500 Hz are shown in the right column (C1-C6). Grand average results across all subjects are shown. It should be noted that the ordinate of panels in the right column is different from that in the middle column.

Figure 2 show the temporal waveforms and amplitude spectra of FFR evoked by the six stimuli with different durations. As stimulus duration increased, the duration of sustained response increased (Figures 2(A1-A6)) and the spectral peaks at harmonic frequencies became more and more prominent (Figures 2(B1-B6) and 2(C1-C6)). Though the same data processing procedures were used for all FFR signals, the spectral resolution could be different since there were different cycles included in FFR signals (Figures 2(A1-A6)). However, the different spectral resolution could not affect the FFR spectral amplitude. To objectively examine the effects of stimulus duration on FFR, its amplitude was measured both in the time domain (FFRrms) and the frequency domain (FFRlowFreqAmp and FFRhighFreqAmp). The FFR latencies of different durations were also compared.

3.2. Effects of stimulus duration on FFR

FFR measures trends with increasing stimulus duration were shown in Figure 3. The amplitude of sustained response FFRrms increased with stimulus duration and the slope tended to be much slower after 100 ms (Figure 3(A)). A two-way ANOVA indicated the change of FFRrms was significant [$F(5,45)=11.43$, $p<0.001$]. Bonferroni corrected multiple comparisons showed that difference emerged for the 50 ms duration, significantly smaller than other durations ($p\leq0.001$) except 75 ms. Also, FFRrms for 75 ms was significantly smaller than 150 ms ($p=0.042$) and 175 ms ($p=0.009$). No significant increase for FFRrms occurred after stimulus duration achieved 100ms. FFR amplitude in

frequency domain was also analyzed, as shown in Figures 3(B) and 3(C). ANOVA analysis showed significant effects of stimulus duration both on FFR amplitude at lower frequencies FFRlowFreqAmp [F(5,45)=2.68, p=0.033] and on FFR amplitude at higher frequencies FFRhighFreqAmp [F(5,45)=2.51, p=0.044]. Multiple comparisons revealed that FFRlowFreqAmp was significantly smaller when evoked by 50 ms stimulus than by 100 ms stimulus (p=0.015), and FFRhighFreqAmp for 50 ms was significantly smaller than at 175 ms (p=0.05). No significant change happened for the latency of sustained response over the durations of 50 ~ 175 ms [F(5,45)=2.23, p=0.057]. Linear regression analyses showed significant correlations between FFRrms amplitude and stimulus duration [R2=0.85, p=0.008], and between FFRhighFreqAmp and stimulus duration [R2=0.86, p=0.007]. No significant linear relationship was found between FFRlowFreqAmp and stimulus duration. Both FFRrms amplitude and FFRlowFreqAmp increased largely from 50 ms to 100 ms, then tended to be saturated with longer durations (125 ms ~ 175 ms), while FFRhighFreqAmp showed a more linear relationship and increased with longer durations (Figure 3). FFR amplitude growth with stimulus duration, as well as the correlative relationships, indicated the existence of ATI at brainstem level.

3.3. Effects of stimulus duration on cortical response P1

The waveforms of auditory late responses (P1 peaks) were shown in Figures 4(A1-A6). The effect of duration on the magnitude of P1 was not consistent with the effect on the amplitude of FFR, for P1 amplitude tended to decrease as increases in stimulus duration (Figure 4(B)). ANOVA analysis

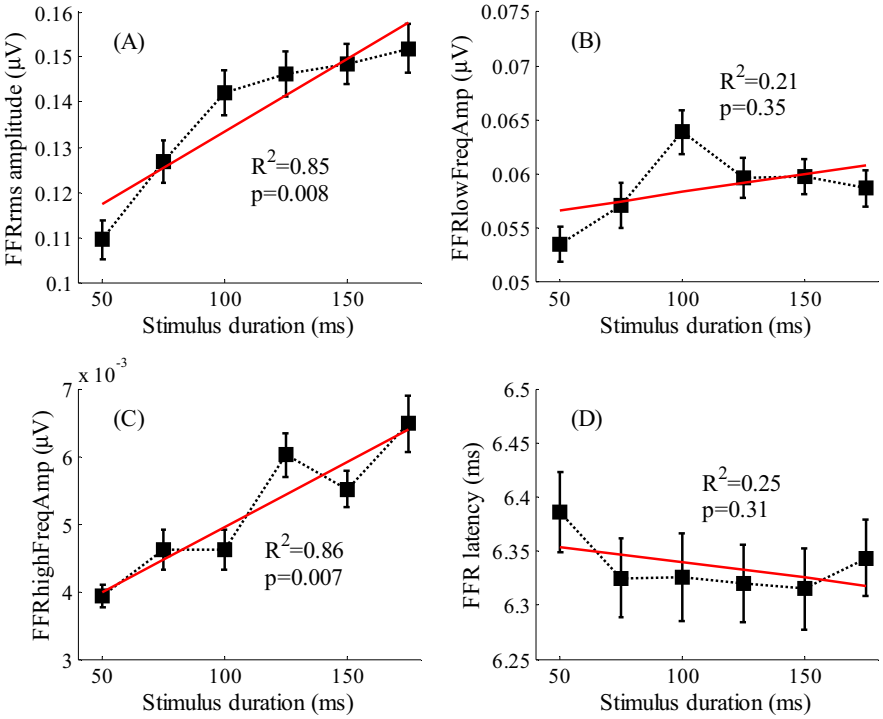

Fig. 3. Effects on FFR measures of stimulus duration. FFR measures include temporal amplitude FFRrms in (A), spectral amplitude at lower frequencies FFRlowFreqAmp in (B) and at higher frequencies FFRhighFreqAmp in (C), and FFR latency in (D). Linear regression result is illustrated with the solid straight line in each panel. Grand averages across all subjects are shown, with the error bar representing one standard error.

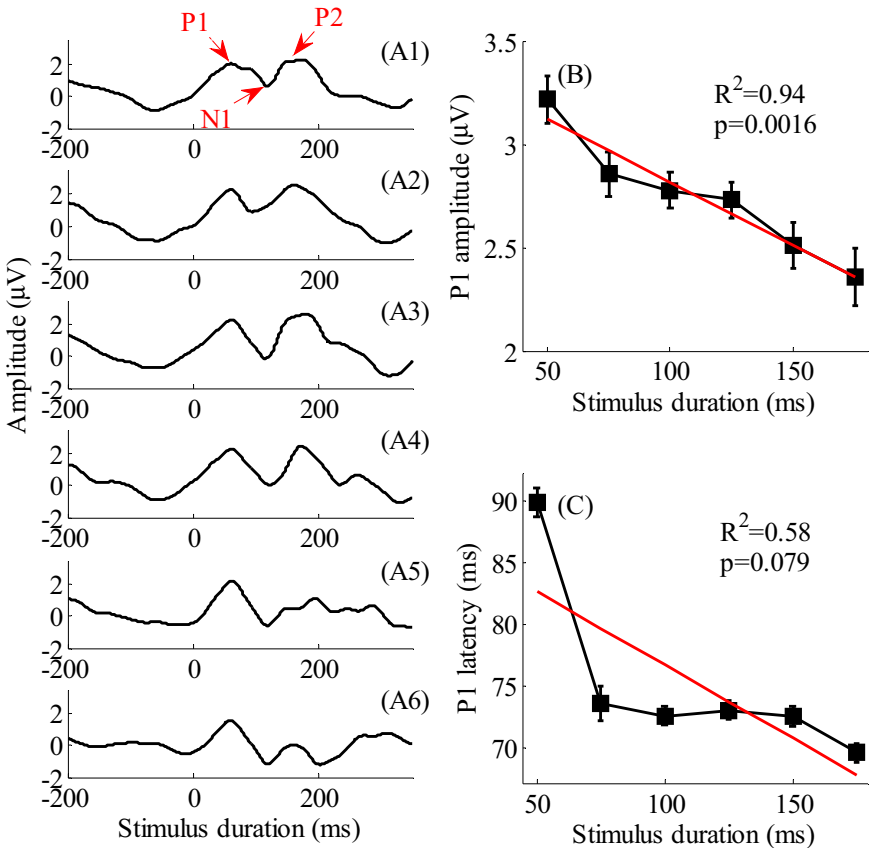

Fig. 4. Cortical response waveforms and effects of stimulus duration on P1 peaks. Panels in the left (A1-A6) show the waveforms of P1-N1-P2 evoked by the six durations, which are 50 ms, 75 ms, 100 ms, 125 ms, 150 ms and 175 ms from top to bottom. Data of one subject are shown in the left panels (A1-A6). The trend of P1 amplitude with increasing durations is shown in Panel (B), and the trend of P1 latency in (C). Other conventions are the same as for Figure 3.

revealed that the P1 amplitude was affected significantly by durations [$F(5,45)=4.29$, $p=0.003$]. Post hoc Bonferroni corrected multiple comparisons showed it to result from the shortest duration 50 ms evoked significantly larger P1 amplitude than the two longest durations did [150 ms: $p=0.016$; 175 ms: $p=0.002$]. The changes in durations also affected the P1 latency [$F(5,45)=16.26$, $p<0.001$], with 50 ms evoking the latest P1, later than all other durations ($p<0.001$). No significant difference occurred for the P1 latencies over durations from 75 ms to 175 ms. Linear regression gave a coefficient of $R2=0.85$ ($p=0.0016$) for the relationship between the P1 amplitude and stimulus duration. Significant effects on P1 peaks of stimulus duration implied a cortical ATI.

3.4. Behavioral loudness trend with stimulus duration

Loudness increased with stimulus duration, showing a tendency of saturation (Figure 5(A)). A two-way ANOVA revealed a significant stimulus duration effect on loudness perception [$F(5,45)=58.04$, $p<0.001$]. The shortest stimulus (50 ms) yielded the softest loudness, significantly different from other durations ($p<0.001$). The second shortest stimulus (75 ms) was judged to be significantly softer than stimuli with longer durations (100 ms: $p=0.002$; 125 ~ 175 ms: $p<0.001$). The

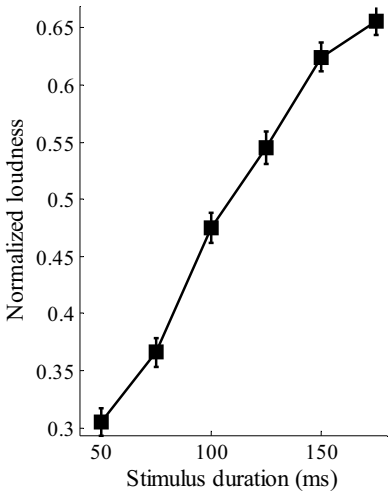

Fig. 5. Behavioral loudness trend with stimulus duration. Grand averages across all subjects are shown, with the error bar representing one standard error.

100 ms stimulus was judged softer than 150 and 175 ms stimuli ($p<0.001$), and the 125 ms stimulus was softer than 175 ms ($p=0.002$). No significant difference occurred between 125 vs. 150 ms and 150 vs. 175 ms. Loudness grew rapidly as stimulus duration increased from 50 to 100 ms, with a significant growth in loudness per 25 ms increase in duration. Then, no significant loudness growth happened for 25 ms increase when stimulus duration was up to 125 ms. That is to say, perceptual loudness grew significantly with increasing stimulus duration up to 100 ms, and then showed a tendency of saturation with a slower growing slope.

3.5. Correlations of electrophysiological and behavioral representations of ATI

Pairwise correlations between electrophysiological and behavioral measures of ATI are shown in Figure 6. Behavioral loudness perception was positively associated with brainstem FFRrms amplitude [$R2=0.89$, $p=0.004$; in Figure 6(A)]. That is, larger FFR amplitude at the level of the brainstem predicted louder behavioral perception. A similarly strong but negative correlation was found between cortical P1 amplitude and behavioral loudness perception [$R2=0.9$, $p=0.004$; in Figure 6(B)]. Significant correspondence was also observed between brainstem FFR amplitude and cortical P1 amplitude [$R2=0.87$, $p=0.007$; in Figure 6(C)]. The strong correlations among brainstem, cortical and behavioral measures implied an interplay between ATIs at different levels along the auditory pathway.

Figure 6(D) illustrates a 3-dimentional scatter with each point in the space representing brainstem, cortical and behavioral measures in the x, y and z dimensions, respectively. A significant correlation between all three measures was revealed by the plane fitted to the data [$R2=0.93$, $p=0.018$], stronger than any single pairwise correlation alone (Figures 6(A-C)). The improvement in the ability to predict behavioral loudness perception using both brainstem and cortical measures reflects coordinated processes with regards to ATI.

4. Discussion

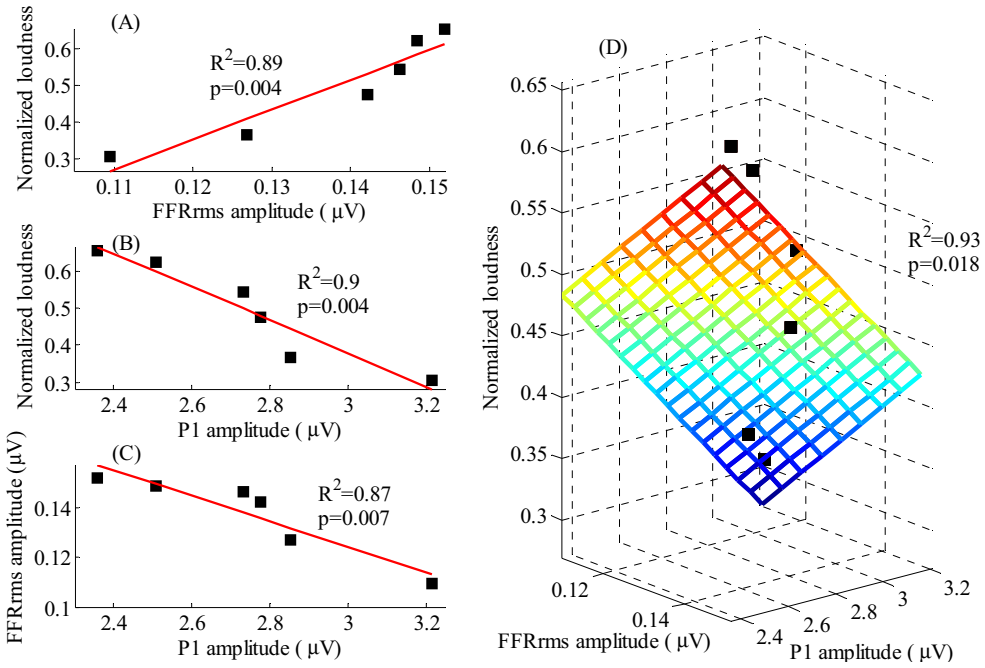

Fig. 6. Correlations of neural and behavioral measures of ATI. Panels in the left show pair-wise regressions among neural and behavioral measures: (A) brainstem FFR amplitude FFRrms vs. behavioral loudness estimated by CMM method; (B) cortical P1 amplitude vs. behavioral loudness; (C) P1 amplitude vs. FFR amplitude. Panel (D) shows simultaneous correlations between all three response measures. Linear regression results are illustrated by the solid straight lines in panels (A-C) and the plane fitted to the data in (D). Each data point represents the grand average across all subjects.

Previous studies have explored the effects of stimulus duration on several AEPs and AEFs, so as to find the electrophysiological correlates of ATI. Among them, later AEP/AEF such as P1-N1-P2 complex was most investigated to reflect ATI at cortical level. Tone or noise bursts were often used as stimuli, and N1 amplitude tended to increase with stimulus duration up to a duration about 18~76 ms [3-7]. The detection threshold for N1-P2 complex decreased with duration increased from 8 to 128 ms, consistent with behavioral threshold change with duration [9]. When consonant-to-vowel syllables (with durations of 120~180 ms including 20~80 ms transitions) were used as stimuli, the amplitudes of N2, P2 and N4 increased with increasing duration for children with language impairment [8]. In our study, the stimuli were presented at a rate faster than twice per second. Instead of N1, the P1 was the most prominent wave of the response and was also the most stable and repeatable across all subjects. We found that P1 amplitude decreased linearly with stimulus duration and was also strongly correlated with behavioral loudness growth. Consistent with previous studies, our results supported the existence of cortical ATI.

In previous studies, to explore ATI at the level of brainstem, ABR evoked by clicks or tone-bursts was usually used since the wave V's origin was the termination of the lateral lemniscus of inferior colliculus [23]. ABR detection threshold as a function of increasing duration was also measured, which was found to be independent of stimulus duration ranging from 1 to 256 ms [12]. In another study [10], ABRs were evoked by pulse series. Comparisons between the ABRs to the first pulse and to the last pulse were performed to assess the effects of ATI and no difference was found when stimulus sound level was around behavioral threshold. These previous studies on ABR suggested that ATI was not found in auditory brainstem. The problem is that ABR was an onset response always

evoked by transient pulses, and it was not affected by the stable ongoing portion of the stimuli. That could be the main factor for its failure to measure ATI at brainstem level. On the contrary, FFR was a better candidate for investigating ATI at the brainstem level because it was a sustained response faithfully following the stimulus. In our study, FFR amplitude progressively grew with increasing stimulus duration and showed a tendency to saturate after 100 ms (Figure 3(A)). FFR amplitude was more strongly correlated with behavioral loudness [R2=0.89, Figure 6(A)] than stimulus duration [R2=0.85, Figure 3(A)]. Our results indicated the existence of ATI at the level of brainstem. Moreover, behavioral loudness was better predicted using both brainstem FFR and cortical P1, suggesting that brainstem ATI was coordinated with cortical ATI (Figure 6).

5. Conclusion

Both brainstem FFR and cortical response P1 was measured to evaluate electrophysiological representations of ATIs in the brainstem and cortex, as well as behavioral loudness temporal integration. Significant effect was reported on FFR amplitude of stimulus duration ranging from 50 ms to 175 ms, indicating an ATI at the level of auditory brainstem. Similarly, ATI in auditory cortex was also reflected by P1 amplitude change with stimulus duration. A strong correlation was demonstrated between brainstem and cortical ATI measures, both of which coordinated with behavioral loudness temporal integration. Behavioral loudness temporal integration was better predicted by using brainstem and cortical response simultaneously than each separately. The close correspondence between neural and behavioral representations of ATI suggested an integrated processing over brainstem, cortical and perceptual levels.

Acknowledgment

This research was supported by the National Natural Science Foundation of China under the grant no. 61271133 and Basic Development Research Key Program of Shenzhen under grant no. JC 201105180808A.

References

[1] B.C.J. Moore, An Introduction to the Psychology of Hearing, 5th ed., Academic Press, San Diego, USA, 2003.
[2] L.E. Marks and M. Florentine, Loudness, Springer, New York, USA, 2010.
[3] S. Onishi and H. Davis, Effects of duration and rise time of tone bursts on evoked V potentials, Journal of the Acoustic Society of America **44** (1968), 582–591.
[4] S.L. Joutsiniemi, R. Hari and V. Vilkman, Cerebral magnetic responses to noise bursts and pauses of different durations, Audiology **28** (1989), 325–333.
[5] N.M. Gage and T.P. Roberts, Temporal integration: Reflections in the M100 of the auditory evoked field, NeuroReport **11** (2000), 2723–2726.
[6] T. Okayasu, T. Nishimura, A. Yamashita, S. Nakagawa, Y. Nagatani, S. Yanai and H. Hosoi, Duration-dependent growth of N1m for speech-modulated bone-conducted ultrasound, Neuroscience Letters **495** (2011), 72–76.
[7] B. Ross, J.S. Snyder, M. Aalto, K.L. McDonald, B.J. Dyson, B. Schneider and C. Alain, Neural encoding of sound duration persists in older adults, NeuroImage **47** (2009), 678–687.
[8] R. Čeponienė, A. Cummings, B. Wulfeck, A. Ballantyne and J. Townsend, Spectral vs. temporal auditory processing in specific language impairment: A developmental ERP study, Brain and Language **110** (2009), 107–120.
[9] A.C. Eddins and J.R. Peterson, Time-intensity trading in the late auditory evoked potential, Journal of Speech, Language,

and Hearing Research **42** (1999), 516–525.
[10] B. Lütkenhöner, Auditory signal detection appears to depend on temporal integration of subthreshold activity in auditory cortex, Brain Research **1385** (2011), 206–216.
[11] A. Rupp, S. Uppenkamp, A. Gutschalk, R. Beucker, R.D. Patterson, T. Dau and M. Scherg, The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans, Hearing Research **174** (2002), 19–31.
[12] M.P. Gorga, A. Beauchaine, J.K. Reiland and D.W. Worthington, The effects of stimulus duration on ABR and behavioral thresholds, Journal of the Acoustic Society of America **76** (1984), 616–619.
[13] P. Heil and H. Neubauer, A unifying basis of auditory thresholds based on temporal summation, Proceedings of the National Academy of Sciences of the United States of America **100** (2003), 6151–6156.
[14] P. Heil, H. Neubauer, M. Brown and D.R. Irvine, Towards a unifying basis of auditory thresholds: distributions of the first-spike latencies of auditory-nerve fibers, Hearing Research **238** (2008), 25–38.
[15] A. Krishnan, Frequency-following response, in: Auditory Evoked Potentials: Basic Principles and Clinical Application, R.F. Burkard, M. Don and J.J. Eggermont, eds., Lippincott Williams & Wilkins, Baltimore, 2007, pp. 313-333.
[16] E. Skoe and N. Kraus, Auditory brain stem response to complex sounds: A tutorial, Ear and Hearing **31** (2010), 302–324.
[17] H. Levitt, Transformed up-down methods in psychoacoustics, Journal of the Acoustic Society of America **49** (1971), 467–477.
[18] M. Epstein and M. Florentine, A test of the equal-loudness-ratio hypothesis using cross modality matching functions, Journal of the Acoustic Society of America **118** (2005), 907–913.
[19] M. Epstein and M. Florentine, Loudness of brief tones measured by magnitude estimation and loudness matching, Journal of the Acoustic Society of America **119** (2006), 1943–1945.
[20] T. Campbell, J.R. Kerlin, C.W. Bishop and L.M. Miller, Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography, Ear and Hearing **33** (2012), 144-150.
[21] D. Wile and E. Balaban, An auditory neural correlate suggests a mechanism underlying holistic pitch perception, PloS one **2** (2007), 369.
[22] S.J. Aiken and T.W. Picton, Envelope and spectral frequency-following responses to vowel sounds, Hearing Research **245** (2008), 35–47.
[23] A.R. Møller, Neural generators for auditory brainstem evoked potentials, in: Auditory Evoked Potentials-Basic Principles and Clinical Applications, R.F. Burkard, J.J. Eggermont and M. Don, eds., Lippincott Williams & Wilkins, Philadelphia, 2007, pp. 336–354.