Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Psychological and Physiological Acoustics
Session 3aPP: Auditory Physiology and Modeling (Poster Session)

3aPP28. Accuracy of synchrony judgment between two pulses: effects of variations in cochlear delay amount

Eriko Aiba*

***Corresponding author's address: Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-8-31, Ikeda, 5638577, Osaka, Japan, aiba.eriko@aist.go.jp**

The cochlear delay shifts the arrival of lower-frequency components of an auditory signal slightly but systematically behind that of higher-frequency components. Therefore, even if all of the components of a complex tone physically begin simultaneously, their temporal relation is not preserved at the cochlear level. In our previous study, the accuracy of synchrony judgment was measured using two types of chirps (compensated and enhanced chirps) and a pulse. The compensated chirp had an increasing frequency pattern to cancel out the cochlear delay. An enhanced chirp had a delay pattern that enhances the assumed cochlear delay. The pulse had a usual cochlear delay at the auditory peripheral. As a result, the accuracy of synchrony judgment was the highest in the pulse and higher in the enhanced chirp than the compensated chirp, implying that there is an asymmetric aspect. The purpose of this study is to investigate how our auditory system processes this asymmetric aspect, and to investigate the extent of the amount of temporal collapse was tolerated. We also measured the accuracy of synchrony judgment using stimuli that reverse the cochlear delay (the higher-frequency components arrive behind the lower-frequency components), or enhance the delay of lower-frequency components up to 8 times.

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4799728]
Received 22 Jan 2013; published 2 Jun 2013

INTRODUCTION

Onset synchrony is widely assumed to be an important cue for perceptual unification of multiple sinusoidal sound as a single complex tone [1]. However, even if all of the components physically begin exactly simultaneously, their temporal relation might not be preserved at the cochlear level because of the "cochlear delays." There is the possibility that this phenomenon has some effect on perceptions of temporal information of sound.

In our previous psychoacoustic studies [2, 3], experiments were performed to measure the accuracy of synchrony judgment using stimuli that controlled the cochlear delays. The results of the experiment showed that the synchrony judgment accuracy was highest for stimuli that evoke an intrinsic cochlear delay (intrinsic delay chirp). Furthermore, the synchrony judgment accuracy was higher for stimuli that evoke the enhanced cochlear delay (enhanced delay chirp) than for stimuli that cancelled out the cochlear delay (compensated delay chirp), which implies that there is an asymmetric aspect of temporal processing in the human auditory system.

The purpose of this study is to investigate how our auditory system processes this asymmetric aspect, and also to investigate the extent of the amount of temporal collapse that has no effects on our perceptions of temporal information of sound.

Cochlear Delay

The cochlear delay was caused by the stiffness of the cochlear Basilar Membrane (BM) gradually decreases from the basal side to the apical side (Fig. 1)[4]. Therefore, the higher components of an input wave excite the basal side, while the lower components excite the apical side. Due to the time delay required for wave transmission, the vibration caused by low-frequency components reaches its associated location later than that caused by high-frequency components. This delay largely occurs for components below 1000 Hz, and the vibration in the lowest-frequency associated location is delayed by about 10 ms relative to the vibration in the high-frequency associated location [5]. As a result, the cochlear delay shifts the arrival of lower-frequency components of an auditory signal slightly but systematically behind that of higher-frequency components.

In previous physiological studies [6, 7, 8], the auditory brainstem response (ABR) was observed, when participants listened to the compensated delay chirp, the enhanced delay chirp and the pulse respectively. As a result, the ABR wave-Vs that evoked by the compensated delay chirp appeared to be the most salient, and the enhanced delay chirp had the gentlest peak. It is generally assumed that the ABR is an electrophysiological event evoked by the onset of an acoustic stimulus. It seems that the ABR wave-V reflects the synchrony of excitation along the entire BM partition. If this physiological saliency directly correlates to the perceptual accuracy of the onset detection, the synchrony judgment could become more accurate by compensating rather than by enhancing the cochlear delay

As mentioned above, however, the synchrony judgment accuracy that evoked by each chirp was decreased in order of intrinsic delay chirps, enhanced delay chirps and compensated delay chirps. There is the possibility that the human auditory system performs some compensation for temporal information of sounds, or otherwise it appears to be more tolerant to the delay of the lower-frequency, i.e., the cochlear delay.

In this study, to investigate the extent of the amount of temporal collapse that has no effects on our perceptions of temporal information of sound, we measured the accuracy of synchrony judgment using stimuli that reverse the cochlear delay (the higher-frequency components arrive behind the lower-frequency components), and, on the other hand, enhance the delay of lower-frequency components up to 8 times.

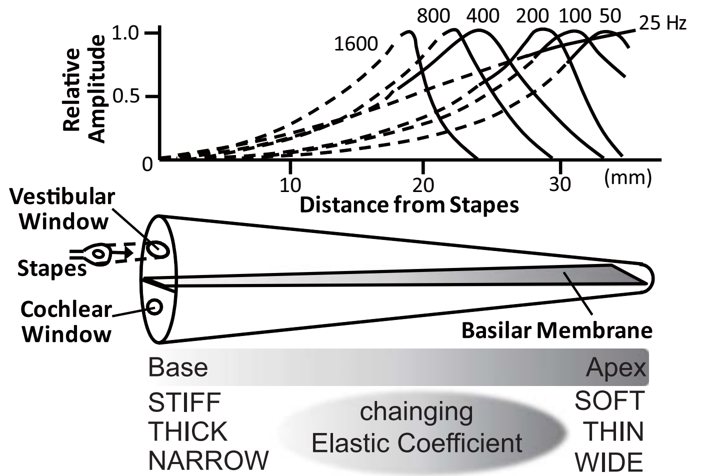

FIGURE 1: The lower part shows cochlea that was straightened and its characteristic features. The Upper part indicates a place of cochlear basilar membrane (BM) where the amplitude of different frequencies reaches the peak respectively.

EXPERIMENT

Estimation of the Accuracy of Synchrony Judgment

To observe the accuracy of synchrony judgment, temporal thresholds to detect synchronous between two pulses were estimated.

Apparatus

The experimental sequences were carried out with a PC (Let's note CF-S10, Panasonic). The stimuli were generated digitally (48 kHz, 24 bit) and controlled by matrix operation software (Matlab, MathWorks) and diotically presented, i.e., the same sound was presented to both ears of participants through insert earphones (IE8, Sennheiser) driven by an audio interface (UA-101, Roland). The participants were seated in a soundproof room and responded to the stimuli by pressing a button on a GUI presented on the PC screen.

Stimuli

Three types of "delay direction" were employed as the experimental stimuli. The second top panel in Fig. 2 shows the (C×1) compensated delay chirp that cancelled out cochlear delay (The broken line in the left panel of Fig. 2). The forth top panel in Fig. 2 shows the (E×1) enhanced delay chirp that enhances cochlear delay. The frequency of the (C×1) chirp increased from 0.1 to 10.4 kHz within 10 ms. In the (E×2) chirp, the frequency decreased from 10.4 to 0.1 kHz within 10 ms. The third top panel shows the (I) intrinsic delay chirp without physical delay imposed on any frequency component was used. The (I) chirp also passed through a low-pass filter with a cut-off frequency of 10.4 kHz.

In addition to the condition of delay direction, we prepared four types of "delay amount" for the (C×1) and the (E×1) chirps. For example, the (C×2) chirp is two times longer than the (C×1) chirp in time to increase the frequency. It means that the frequency of the (C×2) chirp increased from 0.1 to 10.4 kHz within 20 ms. As a result, the (C×2) chirp (The first top panels in Fig. 2) that reverse the cochlear delay (the higher-frequency components arrive behind the lower-frequency components). In a similar way, the (C×4) and the (C×8) chirps were generated.

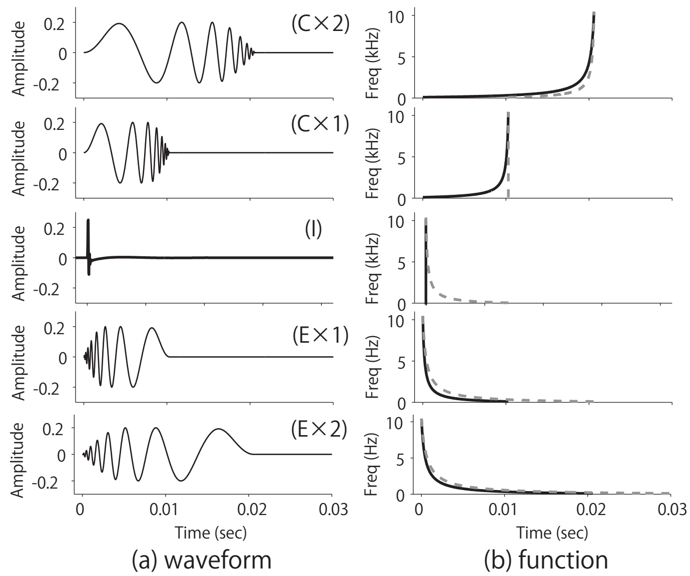

FIGURE 2: The left panels (a) show the waveforms of each chirp. The right panels show the functions of each chirp. The solid line of the right panels (b) shows the frequency pattern as a function of time used under each delay condition. The broken line shows the time required for all frequencies to reach maximum amplitude on the BM.

The (E×2) chirp (The bottom panels in Fig. 2) that evokes two times longer delay than the (E×1) chirp was also prepared. It means that the frequency of the (E×2) chirp decreased from 10.4 to 0.1 kHz within 20 ms, and the (E×4) and the (E×8) chirps were also generated. All chirps except the (I) chirp were tapered transients at both ends with a raised cosine wave of 0.1 kHz.

The level of each stimulus presented was approximately 60 dB SPL. They varied from +6 to -6 dB to prevent participants from applying loudness differences as cues for the experimental task.

Procedure

The experimental task was to detect a synchronous pair in the two-interval, two-alternative forced choice (2I2AFC) procedure. Two pairs of sounds were presented to the participant in one trial (Fig. 3). In each trial, one interval contained a synchronous pair and the other interval contained an asynchronous pair. The asynchronous pairs consisted of 7 variations of a temporal gap (0.2, 0.4, 1.0, 2.3, 5.1, 11.4, or 25.6 ms), spaced roughly logarithmically. The order of the synchronous and asynchronous pairs was randomized across trials to prevent prediction by the participants. The two pairs were separated by a 500-700 ms inter-stimulus interval. The type of the sound was the same within each trial.

Each trial was categorized as one of the following nine types of chirp: I, C×1, C×2, C×4, C×8, E×1, E×2, E×4, or E×8. The total number of stimulus type combinations was 126 (nine types of chirp, seven temporal gaps, and two patterns of order of the synchronous pair). The participants repeated each combination 10 times, which brought the total number of trials to 1260. All factors (sound type, temporal gap, and presentation order) were randomized and executed as a within participant factor.

The participants were instructed that both intervals contained two sounds and that the two sounds were synchronous in one interval but asynchronous in the other interval. They were required to choose the interval containing the synchronous pair. They were able to take breaks at any time. Thresholds were estimated from seven points of psychometric function by fitting a sigmoid function on the results of each participant and computing the temporal gap value

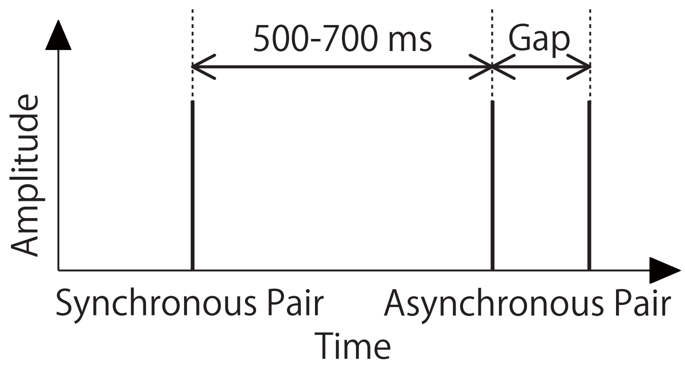

FIGURE 3: Example of stimulus presentation by pulse. In this case, the first interval contains the "synchronous" pair and should be chosen as "correct." Each vertical line represents one of the nine types of chirp.

corresponding to 75% correct responses.

Participants

Seven people (four females and three males) with normal hearing and no history of hearing problems participated.

RESULTS

Figure 4 shows the average estimated thresholds and the standard deviations (SDs) over seven participants for each sound type. A two-way analysis of variance (ANOVA) was performed where the conditions of delay direction and delay amount were treated as the main factors and the participants were treated as a random effect. Delay amount was significant as a main factor ($F(3,18) = 3.84$, $p < .05$). Delay direction was also significant as a main factor ($F(1,6) = 6.58$, $p < .05$). (I) chirp was excluded from this analysis, because it does not have any variations in the condition of delay amount. The interaction between delay amount and delay direction was marginally significant ($F(3,18) = 2.55$, $p < .10$).

To investigate detailed differences in delay direction, we performed one sample t-test for each delay time. Among the (I), the (C×1) and the (E×1), the Tukey-Kramer HSD test was performed. As a result, for the delay amount of one time (×1) and two times (×2), there were significant differences between the delay direction conditions of compensated delay (C) and enhanced delay (E). Thresholds increased as delay amount increased. It means that the accuracy of synchrony judgment decreased as delay amount increased. For delay direction, the thresholds of compensated delay (C) were higher than those of enhanced delay (E). It means that the accuracy of synchrony judgment on compensated delay (C) was lower than that of enhanced delay (E) as previous study reported [3, 9, 7, 8].

DISCUSSION

For delay time, the accuracy of synchrony judgment decreased as delay amount increased. This result suggests that the silent interval between two chirps is one of the important cues to judge onset synchrony. Because of the presence of the shorter duration of a chirp, the longer silent interval arises between two chirps. For example, the duration of the (E×8) chirp is about 80 ms. It means that an asynchronous pair of 25.6 ms gap that consists of two (E×8) chirps has no silent interval. Also, in the case of the (E×4) chirp, there is no silent interval. Actually,

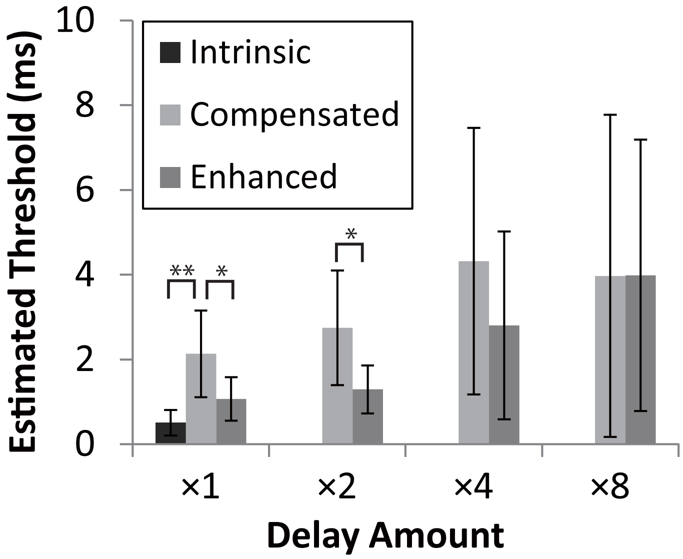

FIGURE 4: The average estimated thresholds in ms and the SDs for the conditions of each sound type ($*p < .05$, $**\,p < .01$).

thresholds for the (E×4) and the (E×8) chirps increase rapidly compared to the (E×1) and the (E×2) chirps. There might be big differences in the task between the (E×4) and the (E×8) chirps, and the (E×1) and the (E×2) chirps.

For delay direction, the accuracy of synchrony judgment on compensated delay (C) was lower than that of enhanced delay (E). It suggests that enhanced delay (E) was perceived as more similar to intrinsic delay (I) than to compensated delay (C). In this study, it was also observed that there is an asymmetric aspect of temporal processing in the human auditory system. The thresholds of the (I), the (E×1) and the (E×2) chirps appear to be almost the same. This is because under the (I), the (E×1) and the (E×2) chirps the lower frequencies reach the associated locations on the BM later than the higher frequencies. This temporal relationship is similar to cochlear delay. In the compensated delay (C), however, the lower frequencies reach the associated locations earlier than the higher frequencies, or otherwise reach almost the same time in the higher frequencies. As we mentioned in Section of Cochlear Delay, the human auditory system appears to be more tolerant to the delay of the lower-frequency than that of the higher-frequency.

As the thresholds of the (E×1) and the (E×2) chirps are almost the same, it is difficult to distinguish the delay of lower frequencies on the BM even if it delays three times more of intrinsic cochlear delay. There is the possibility that the human auditory system is tolerant to the delay of the lower-frequency up to about three times more of cochlear delay.

ACKNOWLEDGMENTS

This work was supported by Grant-in-Aid for the Japan Society for the Promotion of Science (JSPS) Fellows No.24-10633, Grant-in-Aid for Scientific Research (A) No. 24243070 and the JSPS through the "Funding Program for Next Generation World-Leading Researchers (NEXT Program)," initiated by the Council for Science and Technology Policy (CSTP).

References

[1] A. S. Bregman, *Auditory scene analysis* (MIT Press) (1990).

[2] E. Aiba and M. Tsuzaki, “Perceptual judgment in synchronization of two complex tones : Relation to the cochlear delays”, Acoustical science and technology **28**, 357–359 (2007).

[3] E. Aiba, M. Tsuzaki, S. Tanaka, and M. Unoki, “Judgment of perceptual synchrony between two pulses and verification of its relation to cochlear delay by an auditory model”, Japanese Psychological Research **50**, 204–213 (2008).

[4] G. v. Bekesy, *Experiments in hearing* (McGraw-Hill, New York) (1960).

[5] S. Uppenkamp, S. Fobel, and R. D. Patterson, “The effects of temporal asymmetry on the detection and perception of short chirps”, Hearing Research **158**, 71–83 (2001).

[6] T. Dau, O. Wegner, V. Mellert, and B. Kollmeier, “Auditory brainstem responses (abr) with optimized chirp signals compensating basilar membrane dispersion”, J Acoust Soc Am **107**, 1530–1540 (2000).

[7] E. Aiba, T. Shimotomai, K. Kazai, N. Nagata, and M. Tsuzaki, “Judgment of perceptual synchrony between two pulses and its relation to the auditory brainstem response”, Proceedings of the 40th Annual Meeting of Society for Neuroscience (Neuroscience2010) 170.5 (2010).

[8] E. Aiba, K. Kazai, T. Shimotomai, T. Matsui, M. Tsuzaki, and N. Noriko, “Accuracy of synchrony judgment and its relation to the auditory brainstem response: the difference between pianists and non-pianists”, Journal of Advanced Computational Intelligence and Intelligent Informatics **15**, 962–971 (2011).

[9] E. Aiba, “"unification" and "separation" of overlapped sounds.”, Doctral thesis, Graduate School of Kyoto City University of Arts (2009).