Perceptual judgment in synchronization of two complex tones: Relation to the cochlear delays

Eriko Aiba^{1,*} and Minoru Tsuzaki^{2,†}

¹Graduate School of Music, Kyoto City University of Arts,
13-6 Oe-kutsukake-cho, Nishikyo-ku, Kyoto, 610-1197 Japan

²Faculty of Music, Kyoto City University of Arts,
13-6 Oe-kutsukake-cho, Nishikyo-ku, Kyoto, 610-1197 Japan

(Received 10 April 2007, Accepted for publication 16 May 2007)

Keywords: Cochlear delays, Synchronization perception, Complex tones, Perceptual grouping
PACS number: 43.66.Mk [doi:10.1250/ast.28.357]

1. Introduction

The purpose of this study is to investigate whether the delay caused in the course of wave propagation along the basilar membrane of the cochlea (the cochlear delay) has any significant effect on perceptual judgment of synchronization. Onset synchrony of sinusoidal components is widely assumed to be an important cue for their perceptual unification as a single tone. However, even if all the components begin physically in an exactly simultaneous manner, their temporal relation might not be preserved at the cochlear level. This is called the cochlear delay.

1.1. Cochlear Delay

The cochlear delay shifts the arrival of lower components slightly but systematically behind higher components. The cause of the delay is based on the mechanical filtering of the basilar membrane, which works as a transmission line filter bank. In this filter bank, the higher components of an input wave excite the basal end (closer to the oval window), while the lower components excite the apical end. Due to the time delay required for this transmission, the lower components reach the corresponding place later than the higher components. Therefore, even if the components begin simultaneously, their temporal relation is not preserved at the cochlear level.

In the current study, three types of complex tones were used as experimental stimuli to investigate whether cochlear delay imposes a systematic bias in judging the perceptual synchrony of two complex tones.

2. Three types of complex tones

An experiment was conducted to investigate whether cochlear delay affects the detection of asynchrony between two sound objects by using three types of harmonic complex tones: (a) a complex tone with compensatory delay, (b) a complex tone with enhanced delay, and (c) an unprocessed complex tone (with intrinsic cochlear delay). Each complex tone was composed of 30 sinusoids. A temporal relation of 30 sinusoidal components on the compensatory delay condition was altered following the pattern of cochlear delay. The cochlear delay pattern between the components was calcu-

lated on the basis of the model proposed by Dau *et al.* [1]. In this compensatory condition, the delay pattern was designed to “cancel out” the cochlear delay. Accordingly, it can be assumed that excitation along the basilar membrane arises simultaneously so that all the places along the basilar membrane reach the maximum amplitude at the same moment. Figure 1 shows the delay pattern of the compensatory delay condition. The bold line corresponds to the compensatory delay as a function of frequency, and the waveforms of the seven exemplar components out of the 30 are depicted. The enhanced delay condition was generated by using the reversed temporal relation of the compensatory delay condition. Adding to these two experimental conditions, the intrinsic, cochlear delay condition was prepared where no delay was imposed among the components.

3. Experiment

The experimental task was to detect the onset asynchrony between the two complex tones by a two interval, two-alternative forced choice (2I2AFC) task. In each trial, one interval contained a synchronous complex pair, and the other interval contained an asynchronous pair. The order was randomized. Participants were required to judge the synchrony of two complex tones in (a) the compensatory delay condition, (b) the intrinsic cochlear delay condition, and (c) the enhanced delay condition. The time lag between the two complex tones of asynchronous stimuli (delay time) was changed in every trial, and then the thresholds of the three types of complex tones were measured. The delay time was set at 2, 4, 6, 11, 20, 36 and 64 ms using logarithmic scale. Therefore, the thresholds were estimated from 7 points of psychometric function. Figure 2 shows an example of stimuli presentation.

Fundamental frequency and order of starting F_0 were added as random effects. “Synchronous stimuli” has two different complex tones in F_0 and “asynchronous stimuli” also has two complex tones in F_0 corresponding to the synchronous stimuli. The fundamental frequencies of the lower complex tones are 110 or 123 Hz, and the corresponding higher complex tones had an F_0 value higher by an octave. There were two types of starting F_0 orders. The half with asynchronous stimuli starts from the lower complex tones and the other half starts from the higher complex tones.

*e-mail: m06901@kcuu.ac.jp

†e-mail: minoru.tsuzaki@kcuu.ac.jp

Fig. 1 The bold line indicates the delay pattern to compensate for the simulated cochlear delay following the work by Dau *et al.* [1] and the wave forms of a part of the sinusoidal components of the experimental complex tone.

Fig. 2 An example of stimuli presentation. In this case, the second interval contains the “synchronous” pair and should be chosen as the “correct” one.

The number of stimulus type combinations was 168. The participants repeated each of them 10 times. So the total number of trials was 1,680. Feedback was provided after each trial. The level of stimuli was approximately 75 dB SPL and varied +6 to −6 dB.

3.1. Equipment and listeners

The stimuli were generated digitally and controlled by a digital signal processor, Kyma (Symbolic Sound). They were presented to the participants diotically through headphones (SENNHEISER HD 600) connected to an amplifier (Luxman P-1). Participants were seated in a sound-treated room (YAMAHA AVITECS). Four participants with no hearing deficit between 21 and 25 years old participated in the experiment.

4. Results

The ratio of correct choices increased as the delay time increased. It confirmed that participants performed the task using the onset asynchrony for the detection cue as the experimenter intended. Table 1 shows estimated thresholds by all conditions of complex tones. As shown in this table, the threshold of the compensatory delay condition is the highest. An analysis of variance (ANOVA) was performed on this data where the condition of complex tones was treated as the main

Table 1 Estimated thresholds by condition.

Condition	Estimated threshold (ms)
(a) Compensatory	10.8
(b) Intrinsic	7.5
(c) Enhanced	7.7

factor and participants were treated as a random effect. The result of ANOVA was significant [$F(2, 6) = 7.3802$, $p < 0.0241$]. The thresholds of the enhanced delay and the intrinsic cochlear delay conditions are almost the same level.

5. Discussion

The threshold of the compensatory delay condition was the highest and the thresholds of the enhanced delay condition and the intrinsic cochlear delay condition were almost the same level. It suggests the accuracy of the synchrony judgment did not improve even if the cochlear delay was compensated for. On the contrary, such a compensatory operation appeared to make synchrony judgment more difficult. The result that the thresholds of the enhanced delay condition and the intrinsic cochlear delay condition were almost the same level means that the enhanced delay condition is similar to that of the intrinsic cochlear delay condition for the auditory system. As a result of informal listening to these experimental stimuli, it was difficult to discriminate the sound quality of the enhanced delay tones from that of the tones with the intrinsic cochlear delay. In both conditions, the lower components reach the corresponding place later than the higher components.

The observed similarity between the intrinsic cochlear delay condition and the enhanced delay condition was in accordance with our previous experimental result regarding synchrony judgment of two pure tones [2]. The experimental procedure was almost the same as this experiment with complex tones. However, the listeners judged the synchrony of two pure tones with different frequencies instead of the synchrony of two complex tones. The result indicated that it is difficult to discriminate enhanced type stimuli (the higher pure tone preceded the lower one) from intrinsic type stimuli (two pure tones were generated simultaneously).

The tendency for there to be perceptual similarity between the stimulus with the enhanced delay and the simple pulse, i.e., the pulse with the intrinsic cochlear delay, is consistent with the current experimental result. Uppenkamp *et al.* asked their listeners how “pulse-like” their test stimuli sounded using a simple pulse (with the intrinsic delay), an up-chirp (with the compensatory delay) and a down-chirp (with the enhanced delay) [3]. The result indicated that the down-chirp was perceived as more similar to a simple pulse than the up-chirp. This result also shows a certain level of similarity with ours. Uppenkamp *et al.* explained their result on the basis of an additional temporal integration mechanism at a higher stage of auditory processing, which effectively removes phase differences between channels.

Although it cannot be concluded that the result of the current experiment can be explained by exactly the same mechanism as Uppenkamp *et al.*, it is likely that the auditory

system tends to be more “tolerant” with the case where the lower components arrive after the higher components as happens for a simple pulse, as opposed to the reverse case.

6. Conclusion

It cannot be concluded that aligning the excitation peak along the basilar membrane will improve perceptual judgment of synchronization between complex tones. In addition, the auditory system appeared to “ignore” the delay between the components as far as follows the intrinsic, natural direction, i.e., the cochlear delay.

Acknowledgement

This work was supported by a Grant-in-Aid for Scientific

Research A, No. 16200016, Japan Society for the Promotion of Science.

References

- [1] 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).
- [2] E. Aiba and M. Tsuzaki, “Perceptual judgment in synchronization of two complex tones and its relation to the cochlear delays —Dependent of frequency band and level—,” *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 339–340 (2006).
- [3] S. Uppenkamp, S. Fobel and R. D. Patterson, “The effects of temporal asymmetry on the detection and perception of short chirps,” *Hear. Res.*, **158**, 71–83 (2001).