Insensitivity to the coherence of interaural-time-difference modulation across frequency channels

Shigeto Furukawa1,*, Tsuruyo Nishida1,2, Tadahisa Kondo1 and Kazuhiko Kakehi2

1Human Information Science Laboratory, NTT Communication Science Laboratories, NTT Corporation, 3–1 Morinosato-wakamiya, Atsugi, 243–0198 Japan
2Graduate School of Computer and Cognitive Sciences, Chukyo University, 101 Tokodachi, Kaizu-cho, Toyota, 470–0393 Japan

(Received 14 January 2013, Accepted for publication 8 April 2013)

Abstract: The present study examined the dynamic properties of the across-frequency integration mechanism, specifically the extent to which the information about the direction of changes in the interaural-time difference (ITD) is integrated or compared across frequencies. The stimulus was a complex tone consisting of two sinusoidal carriers, one at 400 and the other at 700 Hz. A sinusoidal modulation in the ITD was imposed on one carrier alone or the two carriers simultaneously. The ITD of each carrier was centered at 0 μs, and the modulation started and ended with the zero phase. ITD modulations, when imposed on the two carriers simultaneously, were in-phase or anti-phase between them. Experiment 1 measured the threshold modulation depth for detecting the modulation with an adaptive method. The thresholds were generally lower when both carriers were modulated than when only one was, indicating across-frequency integration of the information about the presence of modulation. The threshold, however, was not significantly different between the in-phase and anti-phase conditions, even when the modulation rate was as low as 1 Hz. Experiment 2 measured the discriminability between in-phase and anti-phase modulations. Modulation depth was fixed at a supra-threshold value (600 μs). The performance varied largely among the listeners, and it was near the chance level for half of listeners even for a 1-Hz rate. The study failed to present compelling evidence that the auditory system is sensitive to the relative phase of ITD modulations for the conditions tested. This suggests that the directional information of even slow (~1 Hz) ITD modulation is not combined effectively across frequencies, at least for the conditions tested.

Keywords: Interaural time difference, Across-frequency integration, Modulation detection, Modulation discrimination, Motion

PACS number: 43.66.Pn [doi:10.1250/ast.34.397]

1. INTRODUCTION

In a realistic environment, it is common that multiple sound sources exist and move in space independently from each other. Mixtures of sounds emanating from individual sources reach our ears, and they are decomposed into frequency components by arrays of bandpass filters, known as the auditory filters, on the basilar membranes [1]. These auditory filters can be regarded as frequency channels. Depending on the frequency content of the sound sources, individual frequency channels convey information about the locations and motions of the sound sources. In order to track the movement of multiple sources, the auditory system should be able to compare information about the direction of source movement across frequency channels.

The interaural time difference (ITD), as well as the interaural level difference (ILD), is a major cue for sound source localization. The movement of a sound source should be accompanied with a change of in the ITD. The present study examined the extent to which the information about the direction of the ITD change can be combined or compared across frequency channels. There are pieces of evidence indicating that ITD information is combined across frequencies. The presence of a signal at another frequency interferes with the detection or lateralization of a static or dynamic ITD on one frequency (known as the binaural interference; e.g., [2–4]). It is known also that the ambiguity of laterality for a sinusoidal stimulus with a certain ITD (known as the phase ambiguity) can be
resolved when additional sinusoidal signals with the same ITD are presented at remote frequencies (e.g., [5,6]). However, only few studies have explored whether and how the directional information about ITD changes is combined across frequencies. One of them is the study by Saberi et al. [4], which showed that the presence of a tone at another frequency interferes with the detection of a linearly changing ITD at one frequency and that, in certain conditions, the interference is greater when the interfering tone changes in ITD in the same direction as the target tone than when it changes in the opposite direction. This result can be taken as evidence that the binaural system is sensitive to the relative direction of ITD changes across frequencies.

The present study approached the binaural system’s sensitivities to relative directions of ITD changes across frequencies using different psychophysical paradigms. Experiments were designed to evaluate the dynamic properties of the mechanism (if any) that is responsible for across-frequency comparison of ITD-change directions.

2. EXPERIMENT 1: DETECTION OF ITD MODULATION IMPOSED ON TWO CARRIERS

This experiment measured thresholds for detecting ITD modulation(s) imposed on one or two carrier frequencies. Carrier frequencies of 400 and 700 Hz were chosen, for the following three reasons: First, the frequency separation is sufficient to ensure that they are processed in predominantly independent peripheral channels (separated by 3.6 in equivalent rectangular bandwidth, (ERB) [7]). Second, the tones can be perceptually segregated by avoiding a consonant frequency ratio. Finally, it is known that the human listeners tend to show the highest ITD sensitivities for a frequency range around 500 Hz.

The modulations, when imposed on the two carriers simultaneously, were either in-phase (two ITDs changed coherently towards the same direction) or anti-phase (the changes were opposite) across the carriers. The results can be interpreted under a conceptual framework that divides the ascending auditory system into three levels of processes (Fig. 1): A level in which the ITD information is processed in separate frequency channels (separated by 3.6 in equivalent rectangular bandwidth, (ERB) [7]); a level in which the outputs from the two channels are integrated (middle column); and a level in which higher-order processes are performed (right column). If the detectability at a given rate is found to depend on the relative phase of the ITD modulations at the two frequencies, then the information about modulation phases must be preserved at the level where the two cues are integrated. More specifically, it would be expected that detectability should be best and worst when the two modulations are in-phase and anti-phase, respectively, with an assumption that, internally, the two modulations can sum up and cancel each other out under these conditions (compare upper and lower rows in Fig. 1). If, on the other hand, the modulation phase information is already lost at or below the level of cue integration, we can expect no effect of the relative phase.

The modulation rate was varied, aiming to find the auditory system’s upper temporal limit for processing the relative phase.

2.1. Methods

Five listeners (three females and two males; age ranging between 31 and 36 years old) with normal hearing (hearing levels of <30 dB at frequencies between 125 and 8,000 Hz) participated in the experiment. One listener, S13, was author TN. All gave written informed consent, which was approved by the Ethics Committee of NTT Communication Science Laboratories. The experiments were performed in a sound-attenuating chamber.

The stimulus consisted of two tone bursts, as the carriers, at frequencies of 400 and 700 Hz. Each burst had a duration of 1,100 ms (a 1,000-ms flat portion flanked with 50-ms raised-cosine ramps), and a sound pressure level of 70 dB. In the “signal interval” of the forced-choice procedure, a sinusoidal ITD modulation was imposed on
either of the carriers alone or on both carriers simultaneously. The condition in which only a 400- or 700-Hz tone carrier was modulated is referred to as the “400-Hz” or “700-Hz” condition, respectively. A burst with ITD modulation was created by imposing sinusoidal phase modulation with opposite modulator phases on identical tone bursts for the left and right ears. Modulation rates of 1, 2, 4, and 8 Hz were tested, and were always common for the two carriers when they were modulated simultaneously. The ITD modulation was imposed on the 1,000-ms-long flat portion of the tone burst only. The modulation was centered at 0 ms, and started and ended at the 0-degree phase. The relative phase of the modulations on the two carriers was either 0 or 180 degrees (referred to as the “in-phase” or “anti-phase” condition, respectively). The starting direction (left or right) of the ITD modulation (with a fixed phase relationship between the carriers) was randomized trial by trial. The “no-signal” interval of the forced choice task contained no ITD modulation, although phase modulations with the same depth as in the signal interval were imposed for individual ears. In this case, the phase of the modulators was identical between ears; hence, the stimuli were diotic.

Due to the cochlear delay along the basilar membrane, the internal representation of our in-phase or anti-phase stimulus was not precisely “in-phase” or “anti-phase,” respectively. The cochlear delay, however, is less than a few milliseconds [8], which is negligibly small when compared with the period of ITD modulation tested in the present study.

The stimuli were digitally synthesized by a personal computer (48,000-Hz sampling frequency), generated with a digital-to-analog converter (Fast Track Pro, M-Audio), amplified with a headphone amplifier, and presented binaurally through earphones (HDA200, Sennheiser).

Listeners’ detection performance was measured by using a three-interval three-alternative forced-choice (3I3AFC) method. The listener was instructed to indicate, by mouse-clicking a button on a computer screen, the interval for which the intracranial image was perceived to move over time and/or to distribute broadly. Feedback was given after each response of the trial. The inter-stimulus interval was 500 ms.

Detection thresholds were estimated by a transformed up-down method [9]. In the first trial of a session, the depth of the ITD modulation (the maximum ITD deviation from 0 μs; ΔITD) in the signal interval was set at 600 μs. In the following trials, ΔITD was decreased or increased by a fixed factor, after two consecutive correct responses or one incorrect response, respectively. The factor was common for the two carriers. Trials were repeated until eight turnpoints were obtained. The threshold of the session was defined as the geographic mean of ΔITDs at the last five turnpoints. Sessions for each condition were repeated two to eight times, depending on listeners.

2.2. Results and Discussion

Figure 2 plots detection thresholds as a function of modulation rate. Each panel represents one listener or the average across the listeners. Error bars represents standard errors. Black and color lines indicate conditions in which one of the two components were modulated and in which both were modulated, respectively (see the key).
conditions) were generally lower than those when the modulation was imposed on either carrier alone (the 400-Hz or 700-Hz condition). This tendency was clearly seen particularly for lower modulation rates \( (p = 0.0271, 0.0153, 0.0680, \) and \( 0.1095 \) for modulation rates of 1, 2, 4, and 8 Hz, respectively; paired \( t \)-test comparing the log-thresholds for the 700-Hz condition and the mean log-thresholds for the in-phase and anti-phase conditions).

Our main interest in the present study was in whether the detection performance depends on the relative phase of ITD modulations across frequencies. We failed to see the dependence. There was no clear difference between thresholds for the in-phase and anti-phase conditions. A two-way within-subject ANOVA on the logarithms of the threshold data indicated no significant main effect of relative phase (i.e., in-phase versus anti-phase conditions; \( F(1,4) = 0.5530, p = 0.4984 \)) while the main effect of modulation rate was significant \( (F(3,12) = 18.74, p < 0.0001) \).

The modulation on the 400-Hz carrier was less detectable than that on the 700-Hz carrier, and therefore might contribute relatively little to the listeners’ detection judgments in the in-phase and anti-phase conditions. This, however, cannot fully account for our failure to observe the relative-phase effect. First, as described above, there was a clear improvement in detection performance (i.e., decrease of threshold) by combining modulations on the two carriers, indicating that the modulations on both carriers had substantial contributions to the performance. Second, the difference between the in-phase and anti-phase conditions was not observable even at the modulation rate \( (\text{e.g., } 1 \text{ Hz}) \) with which detection thresholds were comparable between the 400-Hz and 700-Hz conditions.

### 3. EXPERIMENT 2: DISCRIMINATION OF THE COHERENCE OF ITD MODULATIONS ACROSS CARRIERS

Experiment 1 was conducted under an assumption that the ITDs at separate frequencies add up or cancel each other out to form a single auditory percept. It was expected that in- and anti-phase ITD changes would result in larger and smaller changes in the centroid of an intracranial auditory image, respectively (see Fig. 1). The amount of the change for the anti-phase condition also would be expected to be smaller than that for the cases when only one carrier was modulated. The results of experiment 1, however, showed that the detectability for the anti-phase condition was greater than for the condition with modulations on a single carrier. This indicates that, contrary to the above assumption, the across-frequency integration of ITD information is not performed in a simple additive/subtractive manner.

A possibility is that the integration mechanism is sensitive to the relative phase of the ITD modulation across frequencies and thus, the outputs of the mechanism differed between the in- and anti-phase conditions. It may be that, in our detection experiment, the listeners achieved nearly the same level of detection performance, because they used different internal cues depending on the relative-phase condition. Thus, it would be premature to conclude from the results of experiment 1 that the binaural system is insensitive to the relative phase of ITD modulation across frequencies.

Experiment 2 was conducted to test more directly whether the binaural system is sensitive to the relative phase. Listeners’ discriminability between in- and anti-phase ITD modulations was measured using a forced choice paradigm.

#### 3.1. Methods

Nine listeners with normal hearing participated in the experiment (four females and five males; age ranging between 20 and 43 years old). Listeners \( S_1 \), \( S_2 \), \( S_3 \), and \( S_4 \) participated also in experiment 1 \( (S_1 \), \( S_2 \), \( S_3 \), \( S_4 \), \( S_5 \), and \( S_6 \) respectively). Listeners \( S_7 \), \( S_8 \) was author SF.

Similarly to experiment 1, stimuli were tone bursts with frequencies of 400 and 700 Hz, sound pressure levels of 70 dB, and durations of 1,100 ms (including 50-ms long onset/offset ramps). The ITDs of the bursts were centered at 0 µs. Sinusoidal ITDs, as modulators, were imposed on both carriers, the same as in the in- and anti-phase conditions tested in experiment 1. Cycles of the ITD modulations started from 0 µs, and had ΔITDs of 600 µs. In the signal interval of a 3I3AFC task, the modulators were anti-phase between the carriers. The remaining two intervals, or the non-signal intervals, had in-phase modulations. The listener was instructed to choose interval that differed from the other intervals in any property (e.g., movement pattern, dispersion) in the intracranial image of the stimulus. The order of the signal and the non-signal intervals varied randomly from one trial to another. The modulation rate was 1, 2, 4, or 8 Hz, and was always common for the two carriers. The modulation rates tested varied among listeners. The 1-Hz rate was tested for all the listeners. The initial starting directions of the ITD modulation were randomized for each presentation, so that they could not serve as cues for conducting the modulation-phase discrimination task. For two listeners, a supplementary condition was included to evaluate the sensitivity to the coherence of static ITD changes on the two carriers. There, static ITDs of 600 µs were imposed on the two carriers. In the signal interval for a 3I3AFC task, the directions of the ITDs differed between the carriers (e.g., the left-ear timing was advanced for the 400-Hz carrier, and delayed for the 700-Hz carrier). In the non-signal interval, the directions were the same for the two carriers. The directions of the ITDs were randomized for each
For both conditions with modulated and static ITDs, feedback on the correct answer was presented after each response. For all the listeners except two (S\textsubscript{27} and S\textsubscript{28}), one block of trials consisted of 50 trials, in which one modulation rate or static ITD was tested. Blocks for each condition were repeated 4 to 12 times depending on the listener, yielding 200 to 600 trials per condition. For listeners S\textsubscript{27} and S\textsubscript{28}, one block consisted of 30 trials. Fifteen of the trials had a modulation rate of 1 Hz, and the remaining 15 trials had a modulation of 2, 4, or 8 Hz. The blocks were repeated eight times for each condition with a random order, yielding results from totals of 360 trials (1-Hz rate) and 120 trials (the other rates).

The apparatus was the same as in experiment 1.

3.2. Results and Discussion

Figure 3 shows the proportion of correct responses in the 3AFC task for the 1-Hz modulation rate, which was tested for all nine listeners. The listeners were ordered and indexed according to their performance. The performance varied somewhat across the listeners, but was generally poor: Four (S\textsubscript{21}–S\textsubscript{24}) of the nine listeners exhibited performance that was not significantly different from the chance level (see the 99% confidence intervals, estimated assuming a binominal distribution). Even for the best-performing listener (S\textsubscript{29}), the proportion was 0.62, which corresponds to \(d'\) of 0.95 [10].

It is difficult to explain the observed inter-listener variability and the poor performance in terms of the amount of training. Even listeners with a large number of trials (\(\geq 500\) trials per condition, marked as open symbols in Fig. 3) showed the same tendency as the rest of the listeners. In addition, we did not observe a general training effect, i.e., improvement with increasing number of trials. For listeners with \(\geq 500\) trials per condition, the trials in chronological order were divided into sections along a time course, each section containing 100 trials. Figure 4 plots the proportion of correct listeners as a function of session number. There was no clear trend common across listeners. A within-subject ANOVA on the data for the first five sessions indicated no significant main effect of session number (\(F(4,16) = 0.419, \ p = 0.79\)). Before conducting the ANOVA, the proportions were converted to \(d'\) [10]. It should be pointed out also that listener S\textsubscript{25} was author SF, who showed near-chance level performance despite his intensive experience in psychoacoustical experiments including binaural-hearing tasks.

Figure 5 plots the proportion of correct responses as a function of modulation rate, for listeners for whom 1- to 8-Hz rates were tested. There was marked inter-listener variability in the overall performance and in the effect of modulation rate. Listeners S\textsubscript{25}, S\textsubscript{26}, and S\textsubscript{28} exhibited relatively poor performances overall, and their performance tended to be highest for the slowest modulation rate (1 Hz). For the rest of the listeners (S\textsubscript{27} and S\textsubscript{29}), the overall performance was good (proportion of correct response ranging between 0.48 and 0.90), and the best performance was achieved at rates of 4 Hz (S\textsubscript{27}) and 8 Hz (S\textsubscript{29}). Those two listeners reported that they used the spread of the intracranial image of the sound as a cue for conducting the task across modulation rates. A within-
subject ANOVA indicated no significant main effect of modulation rate \((F(3, 12) = 0.618, p = 0.62)\). Before conducting the ANOVA, the proportions were converted to \(d'\) [10].

One reason for the generally poor discrimination performance in this experiment might be that the range of modulation rate we tested was above the rate for which the binaural system is most sensitive. Even a rate of 1 Hz may be too fast, when considering that in our daily life, it is rare for us to experience sound sources moving widely at a rate of 1 Hz or higher: An ITD modulation with a \(\Delta ITD = 600\mu s\) (closely corresponding to a sound source at a 45-degrees azimuth for frequencies around 500 Hz [11]) and a rate of 1 Hz is equivalent to a sound source moving at a speed of around 180 deg/s. The results of a supplementary test, however, did not support this explanation. For listeners S\(_{6}\) and S\(_{9}\), an extra condition was included to test their sensitivities to the coherence of static ITD changes (see the Method section). If our choice of modulation rate were the major reason for the observed poor performance, we could expect that the performance with static ITD changes would be markedly higher than that with ITD modulation. The results for the extra condition are shown in Fig. 5, labeled as “Static.” The proportion of correct responses for static ITDs was slightly higher than for ITD modulation, but was still quite small \((d's \text{ of } 0.40 \text{ and } 1.39)\).

4. GENERAL DISCUSSION

Experiment 1 showed that the detectability of simultaneous ITD-modulations of separate carriers was greater than that for the individual modulation. This can be taken as evidence for the existence of an across-frequency integration mechanism for the ITD, which is consistent with earlier studies (e.g., [2–6]).

Experiments 1 and 2, however, consistently failed to present compelling evidence for a mechanism that can effectively sum or compare directional information of ITD modulations across frequency channels, although some indications were found for limited conditions and listeners. The results are, again, consistent with the results by Saberi et al. [4]. Saberi et al. showed that the existence of another frequency component interferes with the detection of linear ITD modulation at a target frequency and that in some conditions, the degree of interference is greater when the interferer’s ITD modulation direction is the same as the target’s than when it is the opposite. However, the effect of the coherence in the Saberi et al. study, although significant, was not large, and was observed for a limited number of conditions: For the longest-duration stimuli \((\text{duration of } 2\text{s, change rate of } 200\mu\text{s/s})\), the coherence effect was observed only when the frequency difference between the target and the interferer was small \(<3 \text{ ERBs} [7])\), implying that the coherence effect was at least partially due to an interaction between the target and interferer within single auditory filters that are tuned to the frequencies between the two components.

An explanation for the observed poor performance is that the information about modulation direction is not effectively transmitted to a mechanism, if one exists, for across-frequency integration of ITD information. If this explanation is correct, the present results imply that the across-frequency integration mechanism is located at the processing level beyond the low-level monaural processes and the binaural processes where ITD and ILD information is combined: It has been reported that listeners are sensitive to the coherence of frequency modulations (FMs) across frequencies at least for modulation rates up to 2 to 5 Hz [12]. This indicates that the information about the modulation direction of the inter-spike interval of the auditory nerve as a basis for FM and ITD processing is preserved at the peripheral level for those rates. Furukawa [13] measured the detectability of simultaneous ITD and ILD modulation on a bandpass noise and examined the effect of the phase relationship of the ITD and ILD modulations. He found a significant effect of the relative phase for up to 10 to 20 Hz modulation rates. This indicates that the directional information about ITD modulation is preserved up to the level where ITD and ILD information is integrated.

The above view of the hierarchy of ITD processing is broadly consistent with existing knowledge obtained from physiological studies. In the auditory pathway, information about the ITD and ILD is processed initially by separate nuclei, namely, the medial and lateral superior
olives, respectively [14]. Both of the nuclei have tonotopic structures, indicating frequency-by-frequency processing. Single neurons in the central nucleus of the inferior colliculus (ICc), a higher-level processing, are known to represent both ITD and ILD information simultaneously [15]. The ICc is, again, tonotopic, indicating that the integration of ITD and ILD is performed also in a frequency specific manner. Across-frequency integration is thought to be performed after these processes by nuclei such as the external nucleus of the inferior colliculus [16].

It is possible that the efficiency of the across-frequency integration mechanism largely depends on various factors. In fact, earlier studies suggest that the across-frequency interference or integration of ITD information is influenced by onset asynchrony and harmonicity of the carriers [2,3,6], the effects of which were not explored in the present study. The large inter-listener variability observed in experiment 2 implies the existence of other unknown factors that could influence the efficiency.

ACKNOWLEDGMENTS

Portions of the present study were presented at the 2012 autumn meeting of the Acoustical Society of Japan and at the 36th midwinter meeting of the Association for Research in Otolaryngology (2013).

REFERENCES