THE ROLE OF SPECTRAL CHANGE DETECTORS IN TEMPORAL ORDER JUDGMENT OF TONES

PACS: 43.66.Mk

Kashino, Makio[1]; Okada, Minae[1, 2]

[1]NTT Communication Science Laboratories, NTT Corporation
2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, JAPAN

[2] Graduate School of Integrated Arts and Social Sciences,
Japan Women's University
1-1-1, Nishi-Ikuta, Tama-ku, Kawasaki, Kanagawa 214-8565, JAPAN

E-mail: mkashino@cslab.kecl.ntt.co.jp (MK)
mokada@st.jwu.ac.jp (MO)

ABSTRACT

Human listeners can judge the temporal order of acoustic events quite accurately in certain conditions. We hypothesized that such accuracy is realized by coding temporal order as a single neural code of spectral change at an early stage in the auditory system. To test this hypothesis, we examined whether adaptation to linear frequency glides affects subsequent temporal order judgment of brief tones. It was found that the point of subjective simultaneity between two tones shifted depending on the direction of spectral change of the adaptor, as predicted by the hypothesis that temporal order is coded by neural units that are sensitive to spectral change direction. It was also found that the amount of aftereffect was significantly reduced when adaptor and test tones were presented in different frequency regions or to different ears, suggesting that the relevant neural units are located before across-frequency integration and binaural convergence.

INTRODUCTION

The temporal order judgment of acoustic events is essential for auditory information processing. Speech or melody, for example, consists of a sequence of brief (< 100 ms) phonemes or musical notes. Each phoneme or musical note is realized as a complex acoustic signal whose spectrum changes rapidly within several tens of milliseconds. Thus, the auditory system must analyze the temporal order of spectral components to identify phonemes or musical notes and put them into the proper order to recognize a sentence or a melody.

A number of studies have examined the accuracy of auditory temporal order judgment by human listeners. Obtained threshold values (i.e., onset asynchrony between acoustic events required for correct identification of temporal order) range from 2 ms (Efron, 1973) to 150 ms (Broadbent & Ladefoged, 1959), depending on various factors such as stimulus pattern, expertise of subjects, and experimental procedure (for review, see Eddins & Green, 1995).

Of special interest here is the threshold value of 2 ms. If the timing of each acoustic event is coded separately in the auditory system, such a fine timing difference of two (or more) acoustic events would be lost due to temporal variations of neural firing in the course of neural transmission and processing at several stages from the basilar membrane to the cortex. Moreover, temporal resolution of neural responses decreases as the auditory pathway is ascended. Whereas the auditory nerve exhibits phase-locked responses to envelope modulation at modulation rates above 1 kHz, the majority of neurons in the auditory cortex are unable to signal envelope modulation at modulation rates of much more than 20 Hz (Schreiner & Langner, 1988).

Considering these physiological limitations, it would be reasonable to assume that fine temporal order of two (or more) acoustic events on a millisecond time scale should be transformed into a single non-temporal (i.e., discharge rate) code at a relatively early stage of auditory processing. Once temporal order is coded as a single code, the content of the code would be free from possible temporal variations during neural transmission to the higher levels of the auditory system. Moreover, once temporal order is coded as discharge rate at an early level of the auditory system where temporal resolution of neural response is high enough, stimulus-synchronized neural response would not be necessary to represent relative stimulus timing at higher levels where temporal resolution of neural response is not high enough.

A candidate for such a single rate code of temporal order of two or more acoustic events that occur within a short period is direction of spectral change. Physiological and psychophysical evidence has been presented for neural units whose discharge rate is sensitive to the direction of spectral change. Neurons responding selectively to upward or downward frequency changes have been found at various neural sites from the cochlear nucleus to the auditory cortex (for review, see Palmer, 1995). Psychophysically, it has been shown that selective adaptation to frequency glides elevates the threshold for detecting linear frequency glides (Gardner & Wilson, 1979), and produces an aftereffect in which a tone having a constant frequency presented after the frequency glides is perceived as if its frequency were changing in the opposite direction to the glides (Kayahara, 1998; Kayahara & Sato, 1999). A similar aftereffect has been shown to occur also in the adaptation to the motion of spectral peaks and notches (Shu, Swindale, and Cynader, 1993). It is possible that such neural units sensitive to spectral changes are activated not only by continuous spectral changes, as the previous studies have shown, but also by successive discrete spectral components provided that they are presented in an appropriate temporal scale as in the case of temporal order judgment. This is similar to visual apparent motion, where a succession of discrete dots in different positions in space is perceived as the continuous motion of a single dot.

In the present study, we utilized the selective adaptation paradigm to evaluate the hypothesis that temporal order of successive acoustic events is coded as a direction of spectral change. If temporal order judgment is affected by adaptation to a continuous spectral change in a direction, it would indicate that the same population of neural units is involved in both temporal order judgment and detection of spectral change direction. In order to gain an insight into the sites of the relevant neural units, we also examined whether the temporal order aftereffect exhibits frequency selectivity (Experiment 1) and interaural transfer (Experiment 2). The frequency selectivity or ear selectivity of the aftereffect, if it exists, would provide evidence for rejecting the possibility that the aftereffect is due to response bias and other factors rather than neural adaptation of spectral change detectors.

EXPERIMENT 1: FREQUENCY SELECTIVITY OF AFTEREFFECT

Methods

Stimuli   A test stimulus was a pair of brief (20 ms including 10-ms raised-cosine onset and offset ramps) tones (Fig. 1). The frequency difference (Δf) between two test tones was 1/3 octave. An adaptor consisted of ten linear frequency glides each traversing 1/3 octave in 50 ms (including 10-ms raised-cosine onset and offset ramps) in either an upward or downward direction. The interval between adjacent glides in the adaptor was 10 ms and that between the adaptor and the earlier test tone was 100 ms. In variable test frequency conditions, the center frequency of a pair of test tones was either 500, 707, 1000, 1414, or 2000 Hz, while that of the adaptor was fixed at 1000 Hz. In variable adaptor frequency conditions, the center frequency of the adaptor was either 500, 707, 1000, 1414, or 2000 Hz, while that of a pair of test tones was fixed at 1000 Hz.

Procedure   The method of constant stimuli was used. In each trial, an adaptor was presented first, followed by a pair of test tones. Subjects were asked to judge which tone was presented earlier, low tone or high tone, in a two-alternative forced-choice (2AFC) task. For each trial, onset asynchrony between test tones was selected randomly from nine values: $-4a$, $-3a$, $-2a$, $-a$, 0, a,

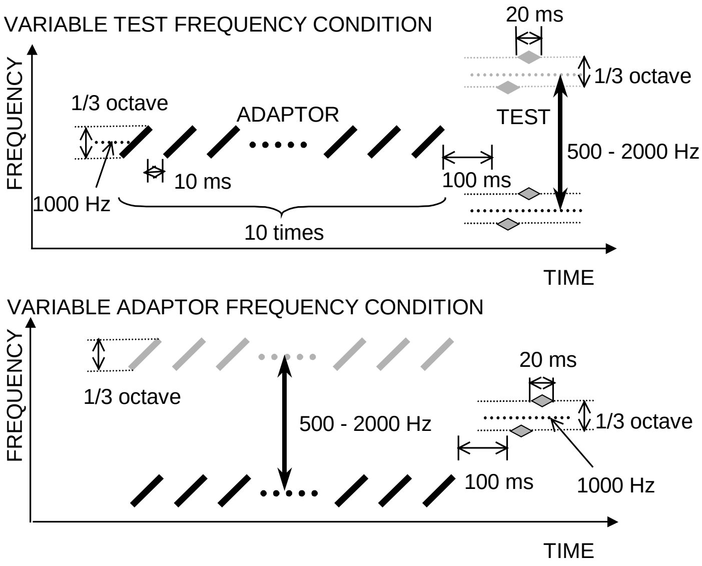

FIG. 1 Schematic representation of the stimuli used in Experiment 1. Upper panel: Stimulus in variable test frequency condition. Lower panel: Stimulus in variable adaptor frequency condition.

2*a*, 3*a*, and 4*a* ms, where step size *a* was either 5, 10, 20, or 50 ms, depending on each subject's ability in temporal order judgment measured in preliminary sessions. Negative values of onset asynchrony indicate that a lower test tone was preceded by a higher test tone. The frequency combination of adaptor and test tones and the direction of frequency change in the adaptor (upward or downward) were fixed throughout a session. Each value of onset asynchrony was tested at least twenty times. In the control condition, only test tones were presented. The time required for a session was 10 - 20 min.

Apparatus All stimuli were generated digitally using MATLAB 5.3 on a computer (Gateway 2000, GP6-400). Sampling frequency was 48 kHz and resolution was 16 bit. The stimuli were presented diotically through a digital audio interface (Digital Audio Labs CardDeluxe), a D/A converter (Stax DAC-TALENT), a stereo amplifier (Nittobo HA-94C), and headphones (Sennheiser, HDA200) at 60 dB SPL. The frequency response of the headphones was compensated by digital filtering. The experiment was carried out in a sound insulated room.

Subjects Five female and two male adults with normal hearing were selected as subjects. They were between 22 and 30 years old. One of the subjects was the author, MO. All seven subjects participated in the variable test frequency conditions, and four of the seven subjects participated in the variable adaptor frequency conditions. All subjects received three hours of training in temporal order judgment of tones in advance of the experimental sessions.

Results

The proportion of responses in which subjects judged that the lower tone preceded the higher was calculated as a function of onset asynchrony between two test tones for each subject in each condition. Then psychometric functions were estimated using the maximum likelihood method.

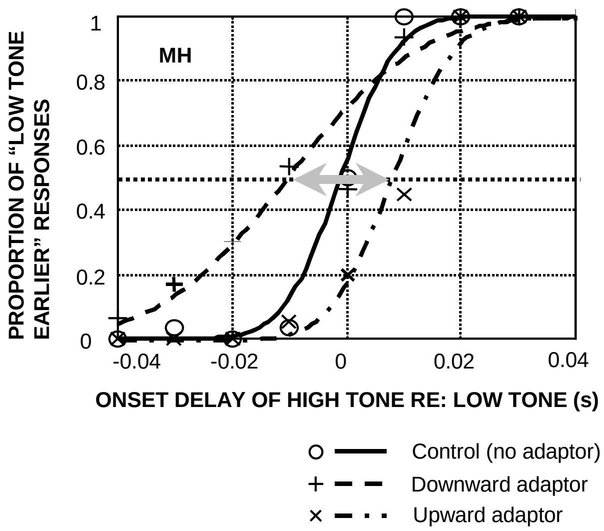

FIG. 2 Examples of psychometric functions obtained from subject MH when center frequencies of adaptor and test tones were 1000 Hz. Symbols indicate raw data. Curves indicate psychometric functions estimated by the maximum likelihood method. Magnitude of aftereffect is defined as the difference of points of subjective simultaneity for upward and downward adaptors.

Fig. 2 shows an example of obtained results. Here the center frequencies of adaptor and test tones were both 1000 Hz. The horizontal axis corresponds to the onset delays of the higher test tone relative to the lower test tone and the vertical axis corresponds to the proportion of "low tone is earlier" responses. The value of onset delay corresponding to a 50% point on a psychometric

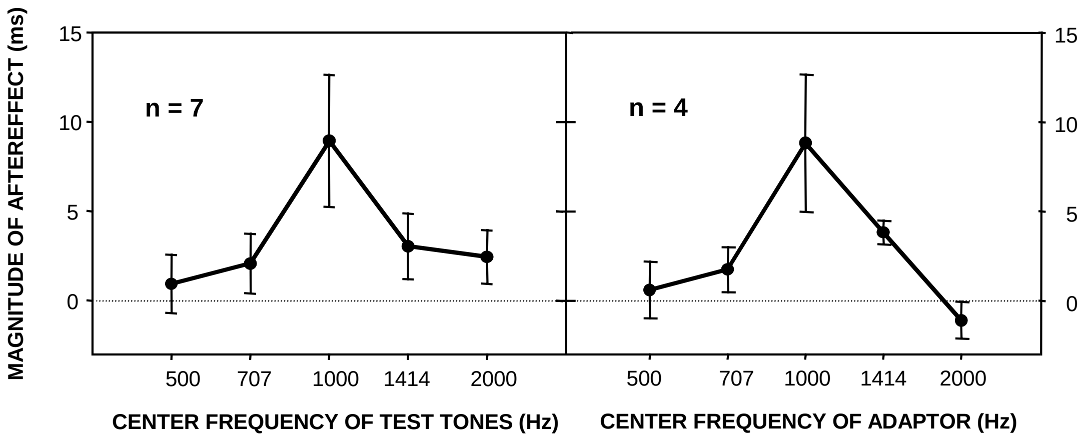

FIG. 3 Frequency selectivity of the aftereffect. Left panel: Mean magnitude of aftereffect as a function of the center frequency of test tones in variable test frequency condition. Right panel: Mean magnitude of aftereffect as a function of the center frequency of adaptor in variable adaptor frequency condition. Error bars show standard errors across subjects.

function is defined as a point of subjective simultaneity for two tones.

The points of subjective simultaneity shifted to a positive direction on the x-axis following adaptation to upward frequency glides and to a negative direction following adaptation to downward frequency glides. This means that two physically simultaneous tones were perceived as if the high tone were earlier than the low tone following adaptation to upward linear frequency glides and as if the low tone were earlier than the high tone following downward linear frequency glides. Here we define the magnitude of aftereffect as the difference between the points of subjective simultaneity in the downward adaptor condition and in the upward adaptor condition.

Fig. 3a shows the mean magnitude of aftereffect as a function of the center frequency of test tones in the variable test tone condition, and Fig. 3b shows that of the variable adaptor condition as a function of the center frequency of the adaptor. In both conditions, the aftereffect decreased significantly when the frequency range of adaptor and test tones did not overlap. The absolute magnitude of aftereffect varied across subjects, but the overall tendency was common for all subjects.

EXPERIMENT 2: INTERAURAL TRANSFER OF AFTEREFFECT

Methods

Stimuli, procedure, apparatus, and subjects were basically similar to those in Experiment 1, except for the following points. First, the center frequencies of adaptor and test tones were fixed at 1000 Hz throughout the experiment. Second, in the same ear condition, both adaptor and test tones were presented to the left ear, whereas in the different ear condition, the adaptor was presented to the right ear and test tones to the left ear.

Results

Fig. 4 shows the mean magnitude of aftereffect in the same ear condition and in the different ear condition. The aftereffect did not disappear completely but reduced significantly when the adaptor and test tones were presented to different ears.

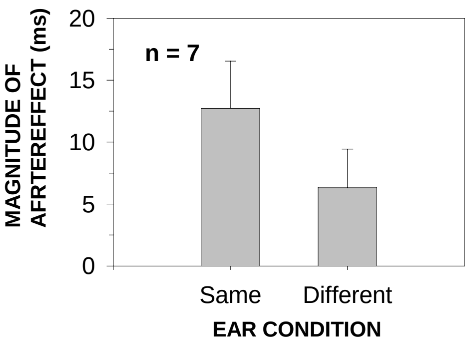

FIG. 4 Interaural transfer of the aftereffect. Mean magnitude of aftereffect is shown for “Same” condition, in which both adaptor and test tones were presented to the same (left) ear, and for “Different” condition, in which adaptor was presented to the right ear and test tones to the left ear. Error bars show standard errors across subjects.

DISCUSSION

We hypothesized that the temporal order of discrete acoustic events that occur within a short period is coded as a single non-temporal code at a relatively early stage in the auditory system. We assumed that such a neural code is provided by neural units whose discharge rate is sensitive to the detection of spectral change. The results of the present experiments are consistent with these hypotheses.

First, the fact that, in certain conditions, adaptation to linear frequency glides *did* affect the subsequent temporal order judgment of brief tones depending on the direction of frequency change of the glides indicates that the same population of neurons is involved both in the detection of frequency change and in the temporal order judgment of discrete tones. The observed shift of the point of subjective simultaneity following adaptation could be explained as follows: According to our hypothesis, a sequence of low and high tones is assumed to activate neural units that are sensitive to upward frequency change, and a sequence of high and low tones activates those sensitive to downward frequency change. When high and low tones are presented simultaneously, the neural units that are sensitive to upward frequency change and those that are sensitive to downward frequency change are activated to the same extent. Now, prolonged exposure to frequency glides in an upward direction would reduce the discharge rate of the neurons that are sensitive to upward frequency change. If two sounds were presented simultaneously in this situation, the neural units sensitive to upward frequency change would be less activated than those sensitive to downward frequency change. This pattern of neural activity is equivalent to that produced by a sequence of high and low tones without adaptation. As a result, temporal order of the simultaneous tones is perceived as if the high tone were coming earlier. This prediction is consistent with the results of the present experiments.

Second, the fact that the aftereffect virtually disappeared when the frequency regions of adaptor and test tones did not overlap indicates that the relevant neural units are located before information from across different frequency regions are integrated. Additionally, the fact that the aftereffect did not disappear completely but reduced significantly when adaptor and test tones were presented to different ears indicates that at least a certain portion of the relevant neural units exists before information from the two ears is converged. These two types of selectivity of the aftereffect are consistent with the hypothesis that the relevant neural site is located at a relatively early stage in the auditory system.

Now let us examine whether any explanations other than our hypothesis are possible for the experimental results.

A common criticism of the selective adaptation paradigm is that the aftereffect does not reflect the sensory process such as reduced response of a certain population of neural units, but is simply due to response bias in decision making (e.g., Wakefield & Viemeister, 1984). Applying this criticism to the present case, the shift of the point of subjective simultaneity of tones following adaptation to frequency glides would be explained as follows: Prolonged exposure to frequency glides in one direction (e.g., low to high) produces response bias in the opposite direction (e.g., high to low), resulting in an apparent negative aftereffect. This explanation, however, would not readily explain why response bias did not affect the results when the adaptor and test tones were presented in different frequency regions or to different ears in the present experiments. Moreover, in the present experiments, the 2AFC paradigm was employed combined with the method of constant stimuli, thus minimizing sequential effect or expectancy effect in decision making. Therefore, it is reasonable to suppose that the results of the present experiments reflect the sensory process rather than response bias in decision making.

Another possible explanation for the aftereffect is based on selective attention; attention drawn to the terminal frequency of the adapting glides facilitates subsequent information processing in the adjacent frequency region, resulting in faster processing of a test tone presented in the frequency region. Support for this explanation comes from the finding that reaction time for the detection of a tone is reduced when attention has been drawn to the frequency region (Schalf, 1998). This explanation would, at least qualitatively, have predicted the direction of the aftereffect observed when adaptor and test tones were presented in the same frequency region in Experiment 1. This explanation fails, however, when frequency regions are different between adaptor and test tones. For example, when the frequency range of the adaptor is higher than that of the test tones, this explanation would predict that the higher tone would be perceived as earlier no matter whether the adaptor is traversing upward or downward, since the higher test tone is closer to the adaptor in frequency than the lower test tone anyway. This is not consistent with the results of Experiment 1. Therefore, we conclude that the aftereffect is not produced by facilitation of information processing elicited by selective attention to the terminal frequency of the adaptor.

We have now excluded two major alternative explanations, and the most plausible explanation for the observed aftereffect seems, at this point, to be the hypothesis we have put forward.

As we have stated in the Introduction, physiological and psychophysical evidence has been presented for neural units whose discharge rate is sensitive to the direction of spectral change. Here we note another recent finding of special interest. Lu, Liang, & Wang (2001) showed that there are two largely distinct populations of neurons in the auditory cortex of awake primates: one with stimulus-synchronized discharges that, with a temporal code, explicitly represent slowly occurring sound sequences and the other with non-stimulus-synchronized discharges that, with a rate code, implicitly represent rapidly occurring events. Furthermore, the neurons of both populations displayed selectivity in their discharge rates to temporal features within a short time-window. Although their experimental settings are not completely compatible with the present study, their finding that rapid time structure is represented as discharge rate of a certain population of cortical neurons is consistent with our idea that fine temporal order of separate acoustic events is transformed into a rate code at an early stage in the auditory pathways and then transferred to higher stages.

The hypothesis that the temporal order of acoustic events is coded as a single rate code of spectral change direction is not only consistent with the current experimental results, but also is applicable to a wider class of perceptual phenomena, such as the effect of duration on the accuracy of temporal order judgment (Eddins & Green, 1995), and various aspects of stream segregation (van Nooden, 1975; Bregman & Rudnicky, 1975; Dorman, Cutting, & Raphael, 1975; Anstis & Saida, 1985; Bregman, Colantonio & Ahad, 1999). It would be a promising avenue of research to examine the relevance of the present theory to those temporal phenomena.

BIBLIOGRAPHICAL REFERENCES

Anstis, S. & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *J. Exp. Psychol.: Human Percept. Perform*., 11, 257-271.

Bregman, A.S, & Rudnicky, A.,I. (1975). Auditory segregation: Stream or streams? *J. Exp. Psychol.: Human Percept. Perform.,* 1 (3), 263-267.

Bregman, A.S, Colantonio, C. & Ahad, P. (1999). Is a common grouping mechanism involved in the phenomena of illusory continuity and stream segregation? *Percept Psychophys.,* 61(2),195-205.

Broadbent, D.E. and Ladefoged, P. (1959). Auditory perception of temporal order. *J. Acoust. Soc. Am.,* 31, 1539-40.

Divenyi, P.L. & Hirsh, I.J. (1974). Identification of temporal order in three-tone sequences. *J. Acoust. Soc. Am.,* 56, 144-151.

Dorman, M.F., Cutting, J.E., & Raphael, L.J. (1975). Perception of temporal order in vowel sequences with and without formant transitions. *J. Exp. Psychol.: Human Percept. Perform.*, 104, 121-129.

Eddins, D.A. & Green, D.M. (1995). Temporal integration and temporal resolution In B.C.J. Moore ed. *Hearing*. San Diego, CA: Academic Press, 207-242.

Efron, R. (1973). Conservation of temporal information by perceptuall systems. *Percept. Psychophys.*, 14, 518-30.

Gardner, R. B. & Wilson, J. P. (1979). Evidence for direction-specific channels in the processing of frequency modulation. *J. Acoust. Soc. Am.,* 66, 704-709.

Kayahara, T. (1998). Changing frequency after-effect of a linear frequency glide. *J. Acoust. Soc. Am.,* 103, 3020.

Kayahara, T. & Sato, T. (1999). Frequency discrimination following linear frequency glide. *J. Acoust. Soc. Am.,* 105, 1152.

Ladefoged, P. & Broadbent, D.E. (1960). Perception of sequence in auditory events. *Quart. J. Exp. Psychol.*, 12, 162-170.

Lu, T., Liang, L., & Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat. Neurosci*,, 4, 1131- 1138.

van Noorden, L.P.A.S. (1975). Temporal coherence in the perception of tone sequences. Unpublished doctoral dissertation, Eindhoven University of Technology.

Palmer, A.R. (1995). Neural signal processing. In B.C.J. Moore ed. *Hearing*. San Diego, CA: Academic Press, 75-121.

Scharf, B. (1998). Auditory attention: The psychoacoustical approach. In H. Pashler, *Attention*. Psychology Press, 75-117.

Schreiner, C.E. & Langner, G. (1988). Coding of temporal patterns in the central auditory system. In G.M. Edelman, W.E. Gall, and W.M. Cowan., *Auditory Function*. New York: Wiley, 337- 361.

Shu, Z.J., Swindale, N.V., & Cynader, M.S. (1993). Spectral motion produces an auditory after-effect. *Nature*, 364, 721-723.

Wakefield, G.H. & Viemeister, N.F. (1984). Selective adaptation to linear frequency-modulated sweeps: evidence for direction-specific FM channels? *J. Acoust. Soc. Am.,* 75(5), 1588-92.