GAP DETECTION IN A TONAL SEQUENCE WITH A FREQUENCY CHANGE

Shuji Mori[1] and Takayuki Sasaki[2]
1. Faculty of Information Science and Electrical Engineering, Kyushu University, mori@brain.is.kyushu-u.ac.jp
2. Miyagi Gakuin Women's University, sasaki@mgu.ac.jp

Abstract

This study investigated how the detection of a gap inserted in a sequence of successive tones would be affected by changing frequencies of the tones. In an experiment, a silent gap was inserted either between the 7th and the 8th tones, the 8th and the 9th tones, or the 9th and the 10th tones in a 12-tone sequence. In the sequence, the frequency of the first 8 tones was always 800 Hz, while the frequency of the following 4 tones was manipulated from 800 Hz to 3200 Hz. The results showed that detection performance was severely impaired when the gap coincided with the frequency change between the 8th and the 9th tones. The impairment was more marked for larger frequency changes and for shorter gap durations. Performance was relatively unimpaired when the gap was inserted before or after the frequency change. These results suggest frequency-uncertainty effects on gap detection (Watson, Foyle, & Kidd, 1990) are due to frequency changes in a sequence, rather than frequency uncertainty per se. The results were also discussed in the context of the perceptual organization of auditory events (Nakajima & Sasaki, 1996; Nakajima, Sasaki, Remijn, & Ueda, 2004; Van Noorden, 1975).

Our auditory system is quite good at detecting a silent pause, or gap, in otherwise continuous sound; a gap of 2 to 3 msec is detectable in broad-band noise and between two clicks (for review, see Scharf & Buus, 1985). However, when a gap is inserted in a sequence of successive tones varying randomly in frequency, its detection requires a much longer duration (Watson, Foyle, & Kidd, 1990). Watson et al. (1990) found that the minimum duration of gap detection corresponding to $d' = 1.0$ (roughly equivalent to 70% of detection probability) was 40 to 50 msec. They attributed the impairment of gap detection to frequency uncertainty in the tonal sequence, where component tones varied randomly in each trial from 300 to 3000 Hz. In our recent observation, gap detection was similarly impaired by simply changing the frequencies of component tones (e.g., from 800 Hz to 3200 Hz), with no frequency uncertainty. Phenomenally, a silent gap was lost between two sub-sequences of different frequencies. In the following, we investigated this 'lost silence' phenomenon in a psychophysical experiment.

Method

Participants

There were 3 participants, 2 graduate students (TK and YS, 24 and 26 years old, respectively) and the first author (SM, 45 years old), all males, and none had difficulty hearing the tones used in this experiment. SM had participated in psychoacoustic experiments for more than 200 hours and YS for about 20 hours before this experiment, while TK had no prior experience of psychoacoustic experiments. YS and TK were naïve to the purpose of this study.

Apparatus and stimuli

Stimulus tones were generated by a sound generation system (Tucker and Davis Technologies, Experimenter II) controlled by a personal computer (IBM, PS/V model 2410), which also controlled timing and data collection. All tones were presented monaurally via a headphone (STAX, SR-Lambda PRO) to the participant’s right ear. Responses were made on a custom-made response box which was connected to the computer via the interface of the sound generation system and was also used to present warning signals and feedback on its LED display.

A tonal sequence was a succession of 12 sinusoidal tones of 75 dB in amplitude and 100 msec in duration, with 3-msec rise and fall times for each tone. There was no silent interval between any of the neighboring tones, other than that introduced as an experimental variable (see below). Of the 12 tones in the sequence, the first 8 tones were always 800 Hz in frequency, while the frequency of the last 4 tones was either 800, 898 (2 semitones above 800 Hz), 1131 (6 semitones or 1/2 octave above 800 Hz), 1600, or 3200 Hz.

A gap was inserted either between the 7th and the 8th tones, the 8th and the 9th tones (where the frequency of the tonal sequence was shifted from 800 Hz to a different frequency for the last 4 tones), or the 9th and the 10th tones. The duration of the gap was set differently for the frequency of the last 4 tones; for 800 and 898 Hz, 5, 10, and 20 msec; for 1131 Hz, 5, 10, 20, and 40 msec; for 1600 and 3200 Hz, 10, 20, and 40 msec. For 1131 Hz, participants SM and TK were first tested with a set of 5-, 10-, and 20-msec gaps, and then an additional experiment was conducted with a set of 10-, 20-, and 40-msec gaps. Participant YS was tested only with a set of 10-, 20-, and 40-msec gaps.

Procedure

Experiments were run in separate sessions for different frequencies of the last 4 tones. During the sessions, the participants first listened to alternative presentations of tonal sequences with and without a gap as often as they wished. They then performed 20-30 practice trials and 90 experimental trials, in which three inserted positions and three durations of gaps were randomly chosen in the stimulus presentations with an equal frequency (10 trials) for each combination of position and duration. Each trial started with the visual presentation of a 500-msec warning signal on the LED display of the response box, followed by successive presentations of tonal sequences with and without a gap. The two sequences were separated by a 500-msec silent interval, and their presentation order was randomly determined in each trial. The participant’s task was two-interval forced-choice (2IFC), that is, indicating which of the two sequences contained the gap by pressing the corresponding button on the response box. There was no time constraint for making a response, and following the response, accuracy feedback was provided on the LED display for 1000 msec. The inter-trial interval was 1000 msec.

The participants first practiced in a session where sequences of 12 tones were all 800 Hz and a 20-msec gap was always inserted between the 8th and the 9th tones in one of the stimulus sequences on each trial. They performed the session with better than 97 % correct responses. They then proceeded to 50 experimental sessions, 10 sessions each for the five frequencies of the last 4 tones, with their order randomized for each participant. The data were collapsed for each participant to calculate the proportions of correct responses (PCRs) separately for the three inserted positions and the three durations of gaps with the five frequencies of the last 4 tones, each based on 100 responses.

.

Results and Discussion

Figure 1 presents separately PCRs for the three participants as a function of inserted position of the gap and frequency of the last 4 tones. Impaired gap detection at the frequency-shifted position is clearly observed. When there was no frequency shifting in the tonal sequence (800 Hz of the last 4 tones), the detection performance was high, even for 5-msec gaps, and mostly independent of the inserted position. When the frequency was shifted from the first 8 tones to the last 4 tones, the detection was impaired at the shifted position (between the 8th and the 9th tones). Impairment was higher for shorter gap durations and larger frequency shift. For 3200 Hz of the last 4 tones, PCRs of the 40-msec gap decreased to about 70%, comparable to the results of Watson et al. (1990) using sequences with frequency uncertainty of the component tones (see Introduction). The results of this experiment suggest that the "uncertainty" effects found were due to frequency changes between successive tones, rather than frequency uncertainty per se.

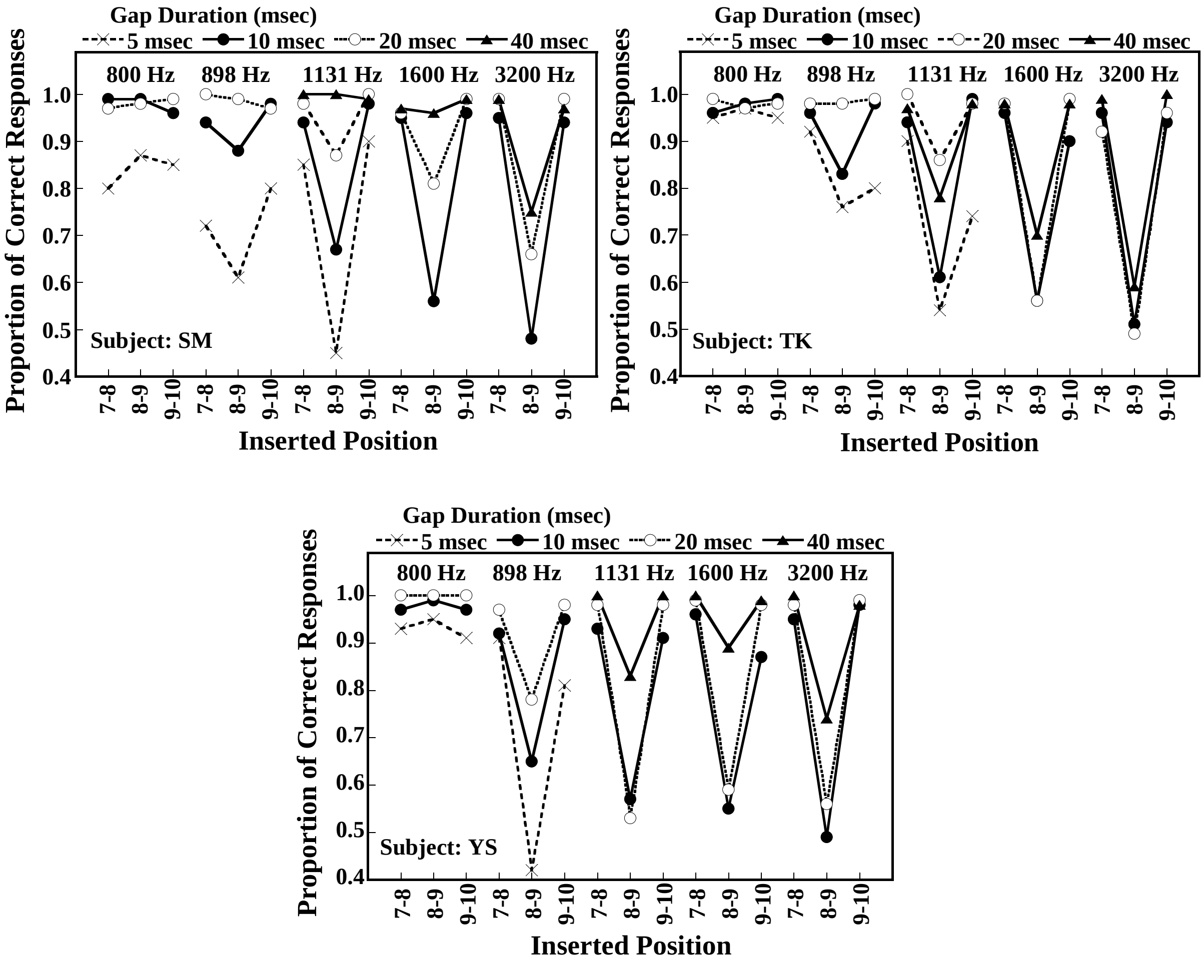

Fig. 1. Proportions of correct responses (PCRs) for each participant as a function of inserted position of the gap and frequency of the last 4 tones.

Detection was relatively unimpaired before or after the frequency-shifted position (between the 7th and 8th tones, the 9th and 10th tones), although the performance was somewhat lowered for large frequency shifts, particularly with short gap durations.

In spite of the clear pattern of results, we have not yet elucidated a definitive explanation for this ‘lost silence’ phenomenon. One possible explanation is an extension of the event construction model of auditory stream formation (Nakajima & Sasaki, 1996; Nakajima, Sasaki, Remijn, & Ueda, 2004). In this model, there are a set of simple grammars by which separate auditory events are formed and segmented into auditory streams. A silent pause (gap) and a change in frequency both mark a stream boundary, if they are salient. They will be redundant as boundary markers when they coincide with each other, and the frequency change will preempt the priority in auditory processing while gap information will receive little processing and eventually vanish. The gap is therefore lost perceptually at the position of frequency change. The phenomenon may also be related to the perceptual ambiguity of temporal relation of separate streams (Van Noorden, 1975). This hypothesis is only speculative, and further research is necessary for this interesting ‘lost silence’ phenomenon.

Acknowledgements

The research reported in this paper was supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science to S. M. (Grant No. 17530528) and to Yoshitaka Nakajima (Grant No. 14101001).

References

Nakajima, Y., & Sasaki, T. (1996). A simple grammar of auditory stream formation [Abstract]. *Journal of the Acoustical Society of America*, *100*, 2681.

Nakajima, Y., Sasaki, T., Remijn, G. B., & Ueda, K. (2004). Perceptual organization of onsets and offsets of sounds. *Journal of Physiological Anthropology and Applied Human Science*, *23*, 345-349.

Scharf, B., & Buus, S. (1986). Audition I: Stimulus, Physiology, Thresholds. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*. New York: Wiley; pp. 14.1-14.60.

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

Watson, C. S., Foyle, D. C., & Kidd, G. R. (1990). Limits of auditory pattern discrimination for patterns with various durations and numbers of components. *Journal of the Acoustical Society of America*, *88*, 2631-2638.