CONFIGURAL PROCESSING IN MELODY RECOGNITION

Jung-Kyong Kim[1] and Daniel J. Levitin[1]
[1]Department of Psychology, McGill University, 1205 Avenue Penfield, Canada, H3A 1B1

1. INTRODUCTION

In the object perception literature, configuration refers to cases in which an object "contains several elements whose spatial relations to each other are perceived as structured and systematic, often giving rise to emergent properties" (Palmer, 1999, p. 707). Most studies of configural processing have been conducted with stimuli in the visual domain; we sought to investigate configural processing in the auditory domain. We chose to study melody recognition because melodies are considered the quintessential Gestalt object, according to Gestalt theory. We hypothesize that a configural processing occurs during melody perception and recognition. By "configural", we mean the processing of an abstract representation of a melody as relying on a specific arrangement of its constituent elements.

Preliminary empirical evidence for configural processing in melodies comes from the detrimental effects of manipulations that disrupt the holistic structure of the melodies, but leave the individual features intact, such as the scrambled melodies employed by Levitin, Menon, and Schmidt (2002). Listeners were to identify well-known melodies that had been divided into small pieces, then scrambled. Listeners reported that the scrambled melodies retained tonality, pitch, timbre, rhythmic, and instrumental information, but they could not recognize the melodies. The authors argued that it is the configural aspects – relations between elements in a melody – that are important for melody recognition.

In the present study, we weakened the influence of featural information of familiar melodies while preserving the configural properties (actual arrangement of the individual elements) of the melodies. Specifically, we created stimuli that minimized local cues to the identity of the melodies by disrupting the pitch sense. This was accomplished by replacing the tones of well-known melodies with bandpass filtered white noise bursts that degraded the pitch quality to such a degree that in the "high noise" conditions, absolute and relative pitch identifications were severely disrupted, while the overall configuration of the melodies remained intact.

We first tested pitch perception by presenting single degraded tones in isolation. Second, we presented pairs of degraded tones. Finally, we presented melodies that consisted of the degraded pitches and tested pitch perception of tones and their overall melody recognition. We refer to these three experimental conditions below as the "contexts" for the pitch identification tasks.

2. METHOD

2.1. Materials & Apparatus

All the stimuli employed were made of filtered noises, with bandpass filters centered symmetrically in log frequency around a centre frequency (F_0). The bandpass filters varied in their bandwidth (or Q: 0.5, 1, 2, 3, 6, 9, 12 semitone-wide).

Single-tone condition included stimuli that were single, filtered noises whose centre frequencies ranged from A4 to G#5 for a total of 12 frequencies. Six different bandwidths were randomly assigned to the 12 different centre frequencies of the filtered noises. Each tone was two seconds in duration. Double-tone or "interval" condition was composed of 12 pairs of filtered noises with the same Q per pair. Six different bandwidths were randomly assigned to the 12 pairs. The centre frequency of the second noise of the each pair ranged from A4 to G#5. The centre frequency of the first noise was randomly assigned from 12 musical intervals (ranging from minor 2nd to unison). The centre frequency of each second noise was randomly assigned to one of the 12 intervals, and the first tone was calculated using the randomly assigned interval size. Each tone was two seconds in duration. For the Melody condition, twenty well-known melodies were chosen based on melodies that are not identifiable by rhythmic cues alone. Each melody was assigned to one of five bandwidths. The centre frequency was assigned such that the last tone of each melody ranged from A4 to G#5. The average duration of the melodies was 14.53 seconds.

2.2. Procedure

Participants were 16 North American adults with at least ten years of music training. They were asked to identify a tone that could range from A4 to G#5 on a piano keyboard. They were allowed to hear the stimuli as many times as necessary, and they could also hum or whistle as an aid to matching the pitch of each tone. For the Single-tone condition, listeners were instructed to identify each single tone, but for the other two conditions, they were to identify the last tone of each pair of tones or each melody. In the Melody condition, participants also provided keywords or a title for each melody.

3. RESULTS

Figure 1 illustrates the average percentages of pitches that were correctly identified for each experimental condition across the different bandwidths (BW). A 3 x7 (Context x BW) repeated measures ANOVA indicated that there was a main effect of bandwidth ($F(6, 90) = 128.94$, $p < .001$), showing that the percentage of correct pitch identification differed significantly across different bandwidths. A main effect of context was not significant ($F(2, 30) = 1.70$, $p = .20$), indicating that there was no difference between contexts in pitch identification accuracy. Also, there was no interaction effect between contexts and bandwidths ($F(12, 180) = 2.02$, $p = .30$).

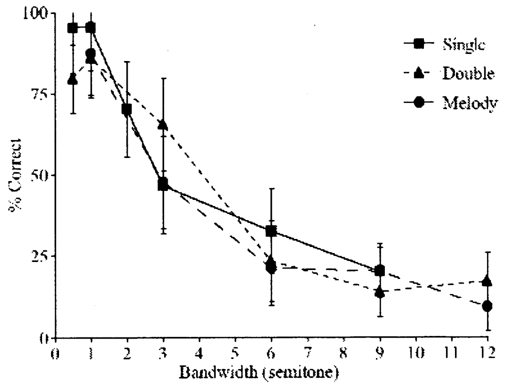

Fig. 1. Percentage of accuracy pitch identification.

We also compared melody recognition and pitch identification for the melody context (Figure 2). There was a significant main effect of task ($F(1, 15) = 137.67$, $p < .001$), indicating that listeners recognized melodies far better than they identified pitches. The main effect of bandwidth was also significant ($F(4, 60) = 34.95$, $p < .001$), indicating that listeners' performance differed across different bandwidths. There was a significant interaction between context and bandwidth ($F(4, 60) = 9.78$, $p < .001$). As shown in Figure 2, the significant interaction suggests that listeners' ability to identify pitches of the melodies declined with decreasing bandwidth while their melody recognition ability did not worsen.

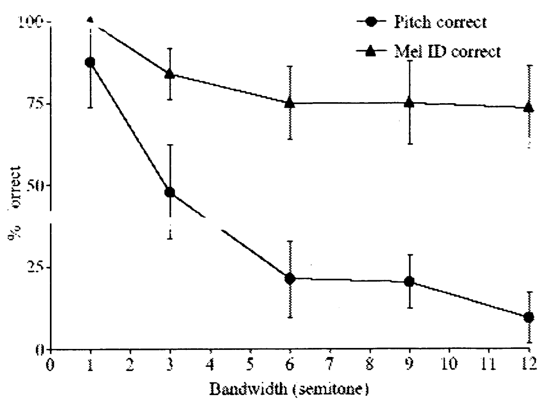

Fig. 2. Percentage of accurate pitch identification and melody recognition.

In order to determine whether listeners had access to contour information, which might have formed the basis for their melody identification judgments, we analyzed their pitch direction (a pitch going up or down) judgments in the Double-tone context. We calculated the percentage of correctly identified pitch directions using only the pitches that were incorrectly identified. The overall percentage of correctly identified pitch directions was 91.5% ($N = 201$), suggesting that listeners generally had a good sense of pitch direction when two tones were presented to them.

4. DISCUSSION

We sought to investigate whether melody recognition was possible on the basis of configural information. Feature-by-feature identification was disrupted since listeners' pitch perception weakened as the bandwidth of the filter increased. We also presented intervals and melodies consisting of filtered noises in order to provide richer contexts. Perception of degraded pitches was not enhanced by adding another tone or a sequence of tones. Most importantly, listeners were able to recognize melodies although they were not able to identify individual pitches that made up the melodies. Their response accuracy was still very high even when the bandwidth of the individual pitches became very wide, thereby disrupting their sense of pitch.

Listeners had a good sense of pitch direction, but this does not imply that melody recognition was based solely contour processing. In previous studies (Dowling & Fujitani,.1970; White, 1960), melody recognition based on contour information was rated much lower than that in our study (59 and 60% repectively). Also, there is evidence (Levitin et al., 2002) that a region in the brain is responsible for the arrangement of musical elements. Taken together, our results suggest that melody recognition does not need to rely on absolute pitch information; even the relative pitch information does not need to be grounded in real and clear pitch elements. Listeners could perceive relations between elements that have no clear identify and use these to access a stored memory trace.

REFERENCES

Dowling, W.J., & Fujitani, D.S. (1970). Contour, interval, and pitch rec ognition in memory for melodies. *The Journal of the Acoustical Society of America, 49,* 524-531.

Levitin, D.J., Menon, V., & Schmidt, J.E. (2002). *The functional neuroanatomy of musical structure* (Tech. Rep. No. 6). Montreal, QC: McGill University, Research Group for Auditory Perception, Cognition, and Expertise.

Palmer, S.E. (1999). *Vision Science: Photons to Phenomenology.* Cambridge, MA: MIT Press.

White, B.W. (1960). Recognition of distorted melodies. *American Journal of psychology, 73,* 100-107.