Temporal Aspects of Melodic Contour Processing

Katherine A. Albrecht

13 June 2014

UC San Diego

ABSTRACT

Melodic contour information is an important cue for song recognition. When listening to novel melodies, contour information tends to be encoded before exact interval information. Alterations to the precise intervals of the melody are often undetected, provided that they do not change the melodic contour. Prior research has suggested that, while the immediate note-to-note contour is most salient, listeners are also sensitive to changes in *non-adjacent* contour information. We were interested in the degree to which this pattern of sensitivity is preserved across musical speeds.

We tested adult listeners' ability to detect changes to novel six-tone melodic sequences with inter-onset-intervals of 50ms, 100ms, 200ms, and 400ms. Changes either preserved contour or disrupted either adjacent or non-adjacent contour.

We found that participants performed at chance in the 50ms condition. Across the other three conditions, the performance pattern was consistent. Adjacent contour violations were best detected, followed by non-adjacent contour violations, with detection of contour preserving alterations generally poor.

INTRODUCTION

Most people can readily identify familiar songs from their culture. Yet no two instances of the same song contain precisely the same acoustic information. The same piece played on two different instruments will differ markedly in timbre. Singers will transpose a piece into different keys based on their individual vocal ranges. Differing familiarity with a particular song can affect the degree of accuracy with which its intervals are reproduced. Yet the ease with which we can judge song identity indicates it must be possible to extract some common feature or features from these varied reproductions.

One feature that tends to be held more-or-less constant across these reproductions is relative frequency. In this study, we will be looking specifically at a simplified version of relative frequency, referred to as melodic contour. This describes possible relationships between pairs of notes in directional terms. Contour can be flat (same note repeated), rising (first note lower in frequency than the subsequent note), or falling (first note higher in frequency than the subsequent note). We are particularly interested in how musical speed affects perception of contour relationships between both adjacent- and non-adjacent tone pairs.

Contour, rather than exact interval information, is often sufficient for song recognition. This is particularly true in the earliest stages of learning. In fact, exact interval information is not accessible to recognition memory immediately following exposure to a novel tune. At two seconds after initial exposure to a 5-note tonal melody, adults are unable to discriminate between its exact transposition and a same-contour lure (Dowling & Fujitani, 1971). As defined in this study, all constituent notes of a tonal melody have membership within a single key of the Western scale system. As delay times increase, false alarms to same-contour lures decrease, provided that the stimuli are tonally strong (Dowling, 1991). When stimuli are either tonally weak or atonal, confusion with same-contour lures persists (Dowling, 1991). In contrast, contour alterations are readily distinguished from exact transpositions, especially when the contour of the original melody is simple (Boltz, Marshburn, Jones, & Johnson, 1985). These early studies focused exclusively on adjacent, note-to-note contour processing. Subsequent work suggests that non-adjacent contour (up to two notes intervening) also contributes to perceptions of melodic similarity, albeit to a lesser degree (Quinn, 1999). However, the subjects in Quinn's study were conservatory students. The extent to which his results apply to the general population is unclear.

Stimuli in cognitive studies of contour processing typically consist of series of 5 to 8 tones, with constituent tones lasting from 175 to 350 milliseconds. Cognitive limitations constrain beat perception to a window of approximately 200 to 1800 ms/beat (London, 2002). The smallest perceptible beat subdivision is around 100 ms (London, 2002). Contour processing at the lowest portion of this range, or at even faster rates, is not well characterized.

Nonetheless, some ERP studies of complex sound processing use an inter-onset-interval of 50 to 100 ms (Schröger, Näätänen, & Paavilainen, 1992; Schröger, 1994; Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen, 2001). All of these studies aim to elicit a mismatch negativity (MMN) component. The first two studies do not explicitly investigate contour. However, in both studies, the better detected of the two deviants violated the adjacent contour of the standard pattern. The better-detected deviant in Schröger's 1992 study violated the adjacent contour of the standard. The other deviant preserved adjacent, but not nonadjacent contour. In Schröger's subsequent 1994 study, both deviants violated both the adjacent and nonadjacent contour of the standard. Tervaniemi et al. explicitly investigated contour (2001). Their five-note standards varied in pitch, but all had an 'inverted U' contour. Deviants were produced by using the same note for both the first and fourth pitch (Tervaniemi et al., 2001). This method resulted in violations to both adjacent and non-adjacent contour. A majority of the subjects had a detection rate below 50%, and only professional musicians completed the task with high accuracy (Tervaniemi et al., 2001).

ERP research requires a large number of presentations of both standard and deviant/target stimuli. Unfortunately, practical constraints limit experiment duration. To the extent that contour processing at rates of 100ms/note or faster is similar to contour processing at slower rates, rapid presentation is a promising solution to this problem.

Currently, it is unclear that this is the case.

The existing literature on auditory streaming suggests that there are likely differences in pitch processing as presentation rate increases. Auditory information is perceptually grouped both by proximity in frequency and proximity in time (Bregman, 1990). At faster rates, proximity in frequency dominates. If the tones in a sequence are sufficiently different in frequency, they will be perceptually divided into high and low frequency streams. Temporal order information is conserved within, but not between, these streams. As melodic contour is a type of temporal order information, it is expected that its encoding will be disrupted at faster rates. The presence and degree of this disruption are likely dependent on both contour complexity and interval size. If all intervals in a stimulus sequence are small, segregation into high and low streams is less likely. Stimuli with simpler contours contain few alternations between low and high notes and therefore are also less likely to induce stream segregation than are stimuli with more complex contours.

The temporal dynamics of auditory object formation may also distinguish contour processing at fast and slow rates. Sequences of auditory events occurring within a window of 150 to 200 ms tend to be processed as a unit (Sussman, Winkler, Ritter, Alho, & Näätänen, 1999). Therefore, the mechanisms for processing successive tones at an inter-onset-interval of 50 or 100 ms may differ from the mechanisms used at an inter-onset-interval of 200 or 400ms. This may have corresponding effects on behavioral performance.

The Present Study

In the presently described study, we tested how the speed of melodic sequences affects detection of contour changes. In particular, we are interested in the relative detectability of changes that alter intervals but preserve contour, changes that alter adjacent contour, and changes that alter non-adjacent contour. We are also interested in changes in overall detection performance across change types.

Using a 2-alternative-forced-choice task, we assessed detection performance in each of these conditions at four different rates of presentation (Experiment 1: 50ms and 200ms;

Experiment 2: 100ms and 400ms). Each subject performed 200 same-different judgments. In each trial, the subject first heard a 6-tone base sequence. After a brief pause, they heard another 6-tone sequence. This sequence could either be identical to the base sequence, or altered at one note. In order to continue to the next trial, the subject had to report whether this sequence was the same as or different from the first sequence.

For longer durations (i.e. 200 and 400 ms per tone), we hypothesized that the results should be consistent with Quinn, 1999. That is, detection of changes that alter any contour information should be better detected than those that preserve contour. Within the subset of changes that alter contour, changes that alter adjacent contour should be better detected than those that alter non-adjacent contour. For shorter durations (i.e. 50 and 100 ms per tone), we hypothesized that the detection of changes that alter contour information should be better than detection of contour preserving changes. However, we predicted comparable performance between adjacent and non-adjacent contour changes. We also predicted an overall decline in accuracy with increasing presentation speed.

EXPERIMENT 1

Methods

Subjects

Participants were recruited from the UC San Diego Human Subject Pool using Sona Systems Participant Management Software. All reported normal hearing.

Experiment 1 included 30 subjects (23 F; age: *M* = 20.7 years; range: 18-24 years). Nine subjects (30%) reported no musical training. Of subjects who reported musical training, mean training duration was 7.76 years (range: 1-17 years; std: 4.96 years).

Stimuli

Stimuli consisted of six-tone sequences. In order to minimize influence from learned scale systems, stimuli were composed of tones from the equal-tempered Bohlen-Pierce scale. The set of possible tone frequencies included the notes in one Bohlen-Pierce octave and the five notes above it (220, 237.6, 261.8, 283.8, 308, 336.6, 367.4, 396, 431.2, 470.8, 512.6, 556.6, 611.6, 660, 719.4, 785.4, 851.4, 924, and 1009.8 Hz).

Stimuli were divided into 4 change conditions: 'same', 'contour preserving', 'adjacent contour violating', and 'non-adjacent contour violating.' The non-adjacent contour violating stimuli had either one or two tones between the two tones whose contour relationship was altered. For all types except 'same', a test sequence was constructed (see Fig. 1 for examples).

Using a PsychoPy script, I pseudorandomly generated 6-unit permutations of this frequency list. The script also generated a corresponding test sequence. Test sequences were identical to the base sequence, with the exception of one note (randomly selected from positions 2 – 5 in the sequence), which was replaced with a random note from the frequency range above. Stimuli were considered valid for inclusion in the experiment if 1) the altered note in the test sequence was absent from the base sequence, 2) the highest and lowest notes of the sequence were identical between the test and base sequence, and 3) only one change type —either adjacent or non-adjacent— resulted from the tone alteration.

Ten base sequences were selected for use in the same, contour preserving, adjacent contour violating, and each of the non-adjacent contour violating conditions. In total, 50 base sequences were produced. For all conditions except 'same,' a corresponding test sequence was produced. All conditions included test sequences with changes at positions 2, 3, 4, and 5. The proportion of changes at each position, magnitude of change between standard and test tones, and contour complexity were comparable between conditions (respectively: $p > .99$, $p > .56$, $p > .66$).

For each stimulus, sound files at rates of 50ms/tone ('short' sequences) and 200ms/tone ('long' sequences) were produced. Stimuli were created using Audacity 2.0.5 sound editing software. Tones were sine waves with on- and off-ramp duration of 5ms (short condition) or 10ms (long condition). Each tone sequence was exported to a .wav file with a sampling frequency of 44.1 kHz.

During the experiment, stimuli were presented through Sennheiser HD 280 headphones. At the beginning of the practice phase, subjects were provided with the opportunity to play a 50ms/tone sample sequence (unused in the main experiment) repeatedly. They were instructed to adjust the output volume until the sequence was at a comfortable level. They were instructed not to change the volume further after the beginning of the practice task.

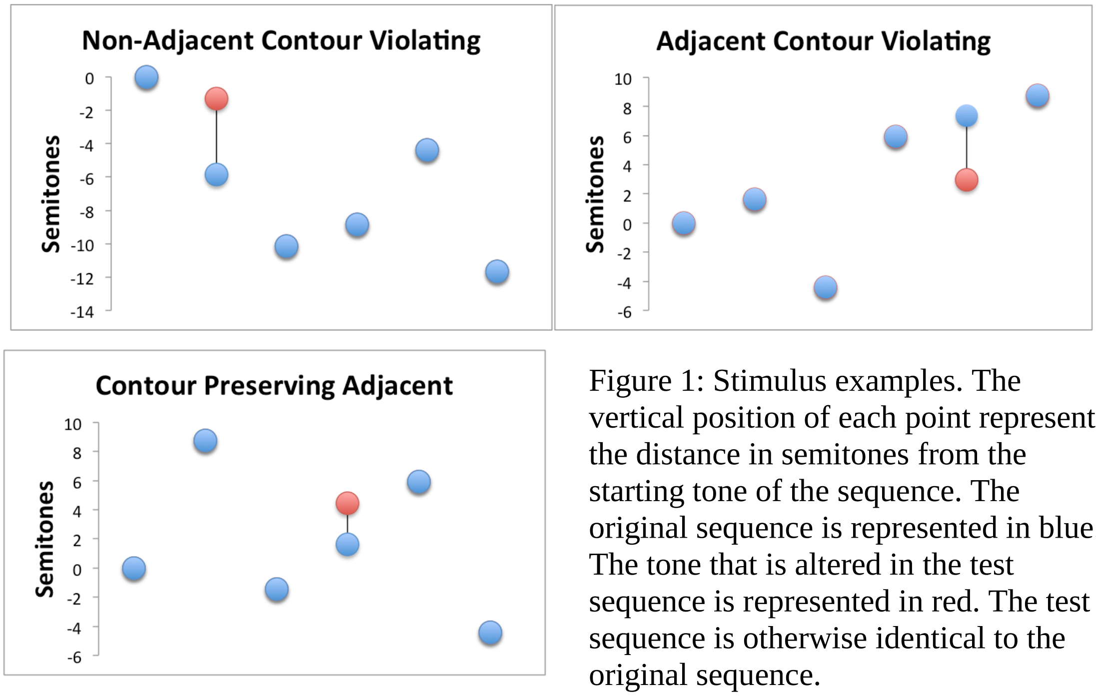

Figure 1: Stimulus examples. The vertical position of each point represents the distance in semitones from the starting tone of the sequence. The original sequence is represented in blue. The tone that is altered in the test sequence is represented in red. The test sequence is otherwise identical to the original sequence.

Procedure

Practice

Upon entering the experiment room, participants completed two surveys. The first survey asked for basic demographic information. The second survey asked participants about their music experience. Survey questions assessed academic music-related coursework, musical training, and listening habits.

The experiment began with a practice task intended to familiarize participants with the procedure. Stimulus display and response collection were handled using PsychToolbox-3 (Brainard, 1997; Pelli, 1997). The practice task began with a demonstration of a short duration 'same' sequence pair. First, a 4-tone base sequence was played. After a 1500ms pause, the sequence was re-played. The word 'same' was displayed in the center of the screen for the duration of this demo. Next, participants heard a demo of a short duration 'different' sequence pair. Once again, a base sequence was played, followed by a 1500ms silent duration. After this gap, an altered version of the base sequence was played. The word 'different' was displayed in the center of the screen for the duration of this demo. Following the demo, subjects were instructed to press 'c' to continue to the practice task.

In the first section of the practice, participants made same/different judgments about 50ms/tone 4-tone stimuli. The interstimulus interval between the base and test sequences was 1900ms. After the test sequence finished playing, subjects were instructed to press the 's' key if they perceived the two sequences as the same, or the 'd' key if they perceived them as different. They had an unlimited amount of time to respond. If the subject gave the correct answer, a green box appeared in the center of the screen with the text “Right!” If their answer was incorrect, a red box appeared in the same location with the text “Try another!” In both cases, participants were given the option to hear the base and test sequences once more.

After completing 5 practice trials at the 50ms/tone speed, the practice task was repeated at the 200ms/tone speed. The demo procedure discussed above was repeated with 4-tone 200ms stimuli. After listening to the demo stimuli, subjects completed 5 same/different practice trials with the long duration stimuli. The procedure was otherwise identical to that of the first practice task. At the end of these five trials, subjects were instructed to press the 'c' key to continue on to the main experiment.

Main Experiment

The experiment was divided into eight blocks, with 27 trials each. Each block contained 5 'same' trials, 5 contour preserving trials, 5 adjacent contour violating trials, and 10 non-adjacent contour violating trials. Two four-tone sequences from the practice condition were also included in each block. All stimuli in a single block belonged to the same duration condition. Duration condition alternated from one block to the next. The four short duration blocks were 1, 3, 5, and 7. The four long duration blocks were 2, 4, 6, and 8. A fixed pseudorandom presentation order was created for each duration. No more than 2 successive trials could be of the same type. No individual stimulus could be used more than once within the same block. In order to control for presentation order effects, half the subjects were assigned to hear the stimuli in the originally generated order (condition A). In order to control for possible fatigue effects late in the experiment, the other half were assigned to hear the stimuli in the reverse order (condition B).

As in the practice condition, stimulus presentation and response collection were handled using PsychToolbox-3. The procedure itself was very similar to the practice procedure. Subjects first heard the base stimulus. After a brief silent interval (1900ms in the short condition, 1000ms in the long condition), the test stimulus played. Instructions to press 's' if the two were the same and 'd' if they were different remained at the top of the screen for the duration of the trial. When the test stimulus finished playing, the text 'Respond now' appeared in the center of the screen. Subjects had an unlimited amount of time to provide their response. After 27 trials, subjects were instructed to take a brief break and press 'c' when they were ready to proceed. At the end of the eighth block, subjects were instructed to press any key to exit the experiment.

Analysis

For each subject, raw hit rates (range: 0 - 1) were calculated for each stimulus type at both durations. Responses to the two four-tone stimuli present in each block were not analyzed. During analysis, errors were found in four stimuli files used in the non-adjacent contour violating condition. Three of these errors were present in both duration conditions. The fourth error was only present in the Short condition. Responses to these stimuli were excluded from the analysis.

In order to correct for guessing, the hit rates for contour preserving, adjacent contour violating, and non-adjacent contour violating stimuli were converted into d' scores. False alarm rates were calculated for each subject at both durations by subtracting the accuracy rate on Same trials from 1. False alarm or hit rates of 0 or 1 were adjusted to 0.001 and 0.999, respectively. Z-scores for both hits and false alarms were calculated using the norminv function in MATLAB. D-prime scores for a given condition were calculated by subtracting the z-score of the subject's false alarm rate from the z-score of their hit rate in the current condition.

Results

A one-way analysis of variance showed no effect of presentation order ($F(1,179) = 0.71$, $p = .402$). Results from groups A and B will therefore be shown together. Since d' scores are comparable for both types of non-adjacent contour violating stimuli in both Short and Long conditions (Short: $p > .30$; Long: $p > .75$), d' results for non-adjacent contour violating stimuli are reported as subjects' average d' over the two non-adjacent conditions.

A two-factor within-subjects analysis of variance showed significant main effects of Change Type ($F(2,179) = 10.05$, $p < .001$) and Duration ($F(1,179) = 65.83$, $p < .001$). The interaction between Change Type and Duration was significant ($F(2,179) = 4.49$, $p < .05$).

Duration

Accuracy on Same trials was comparable between durations (Short: $M = 89.67\% \pm 1.68\%$; Long: $M = 88.83\% \pm 1.71\%$, $p > 0.61$). Accuracy for Different trials was significantly worse for the short duration than the long duration for all stimulus types ($p < .001$ for all change conditions; also see Fig. 2).

Change Type

For the short condition, d' did not exceed chance for any of the trial types (Table 1). In the long condition, d' was above chance for all trial types. Performance information is provided in Table 1. Detection in the adjacent contour violating condition was significantly better than in the contour preserving condition ($p < .001$) and the non-adjacent contour violating condition ($p < .05$). Detection in the non-adjacent contour violating condition was significantly better than in the contour preserving condition ($p < .01$).

Table 1

Type	**Mean d' Short**	**Mean d' Long**
Contour Preserving	-0.11±0.17	0.55±0.20
Adjacent Contour Violating	0.16±0.17	1.90±0.22
Non-Adjacent Contour Violating	0.08±0.15	1.27±0.15

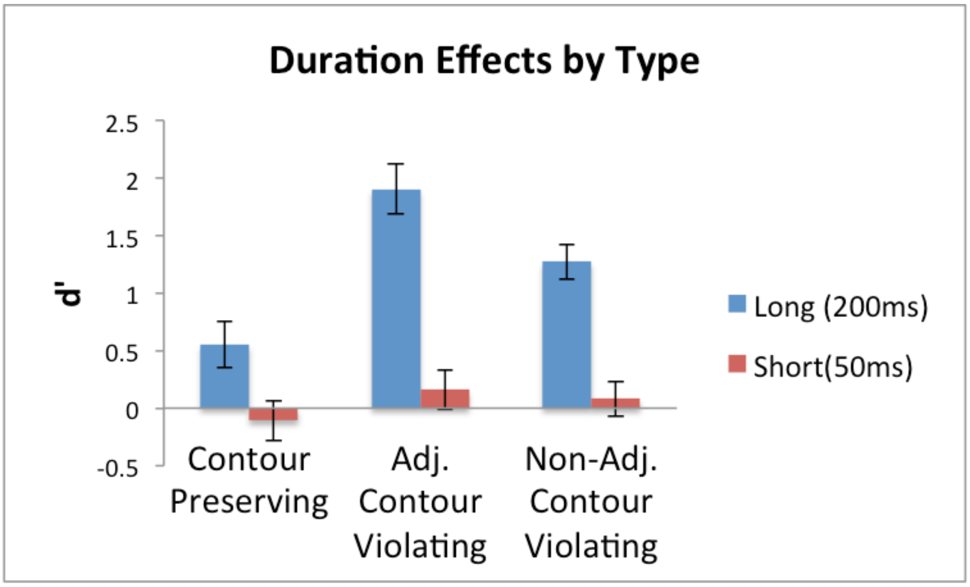

Figure 2

EXPERIMENT 2

Methods

Subjects

Experiment 2 included 33 subjects (21 F; age: *M* = 20.6 years; range: 18-24 years). Eleven subjects (33.3%) reported no musical training. Musically trained subjects reported a mean of 6.86 years of experience (range: 1-16 years; std: 5.34 years).

Stimuli

Except for duration, tone sequences were identical to those in Experiment 1. In the short condition, each tone lasted 100ms. In the long condition, each tone lasted 400 milliseconds. Sound files were exported to .wav format in MATLAB, at a sampling rate of 44 kHz.

Procedure

Subjects performed the same series of tasks as in Experiment 1. The practice condition began with a 100ms/tone demo, followed by 5 100ms/tone practice trials. This was followed by a 400ms/tone demo and 5 400ms/tone practice trials. In the main experiment, all subjects heard short stimuli in blocks 1, 3, 5, and 7 and long stimuli in blocks 2, 4, 6, and 8.

The trial order for the long condition was the same as for the short condition in Experiment 1, and vice versa. Subjects in Condition B heard trials in the reverse order.

Analysis

Procedures were the same as in Experiment 1, except that corrected versions of the previously excluded stimuli were used in this experiment and included in the analyses.

Results

A one-way analysis of variance showed no effect of presentation order ($F(1,197) = 0.05$, $p = .823$). Results from groups A and B will therefore be shown together. A two-factor within-subjects analysis of variance revealed significant main effects of Change Type ($F(2,197) = 13.54$, $p < .001$) and Duration ($F(1,197) = 55.27$, $p < .001$). The interaction between Change Type and Duration was not significant ($F(2,197) = 1.07$, $p = .34$).
Duration: Accuracy on same trials was significantly worse for short than long stimuli (Short: $M = 84.24\% \pm 1.97\%$; Long: $M = 89.39\% \pm 1.61\%$; $p < .05$.) After using d' to control for individual subjects' false alarm rates, accuracy remained significantly worse for the short duration across all change conditions (Fig. 3).

Change Type

Short: D' was near chance in the contour preserving condition, but above chance for both adjacent contour violating and non-adjacent contour violating conditions (Table 2). Detection in adjacent contour violating trials was significantly better than in either contour preserving or non-adjacent contour violating trials (p < .001, p < .01, respectively). Detection in non-adjacent contour violating trials was significantly better than for contour preserving trials (p < .05).

Long: D' was above chance in all conditions. Detection in adjacent contour violating trials was significantly better than in contour preserving trials (p < .001) or non-adjacent contour violating trials (p < .001). Detection in non-adjacent contour violating trials was significantly better than in contour preserving trials (p < .001)

Table 2

Type	Mean d' Short	Mean d' Long
Contour Preserving	0.18±0.16	1.09±0.17
Adjacent Contour Violating	0.85±0.17	2.20±0.23
Non-Adjacent Contour Violating	0.47±0.13	1.38±0.17

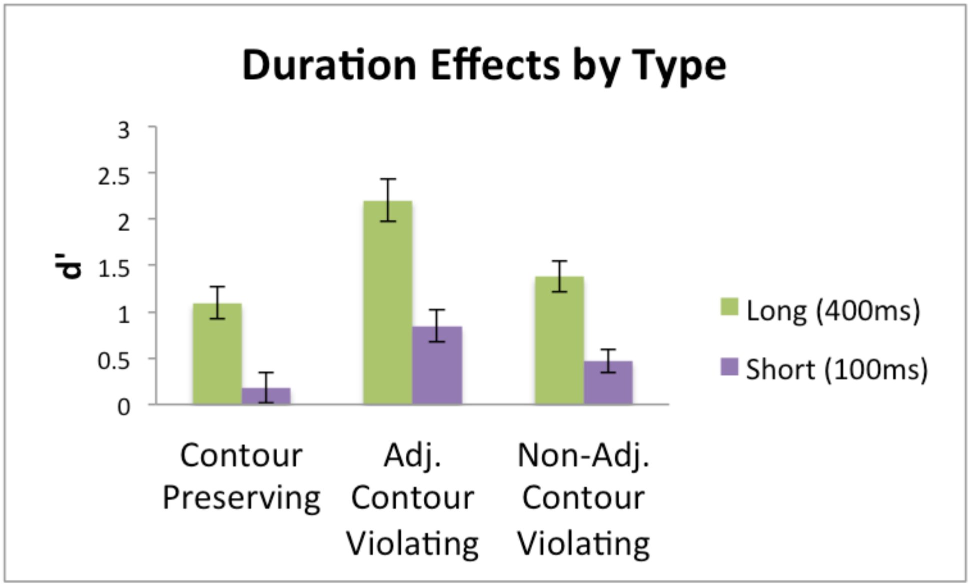

Figure 3

Between-Experiment Comparison

In the 100ms condition, detection improved significantly only for the adjacent contour violating condition relative to the 50ms condition (p < .01). The improvement in performance in the non-adjacent contour violating condition was marginally significant (p = .054). In the 200ms condition, detection improved significantly relative to the 100ms condition for both adjacent contour violating and non-adjacent contour violating conditions (both p < .001). In the 400ms condition, the only significant difference was an improvement in detection of contour preserving changes relative to the 200ms condition (p < .05). Results are illustrated in Figure 4.

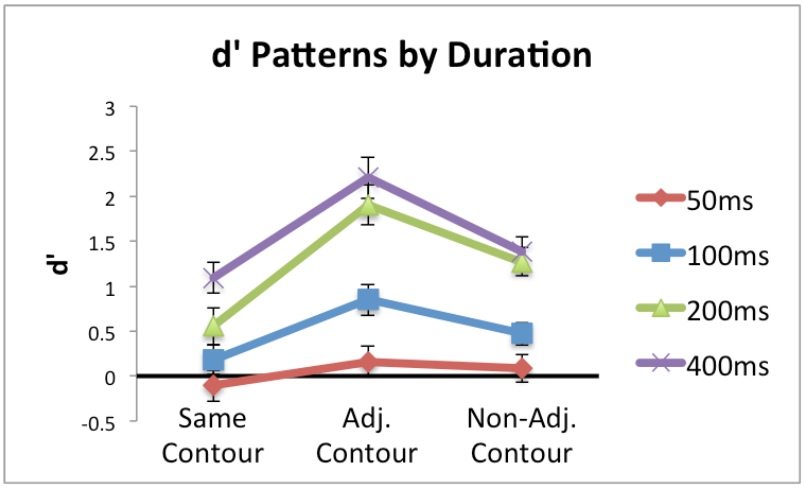

Figure 4

DISCUSSION

The results support our second hypothesis. Overall performance improves as presentation rate slows. This is most obvious in the 50ms condition, where performance in all change conditions was at floor. However, this improvement is not constant with decreasing rate. Nor is the degree of improvement even across change types. Collapsing across change conditions, performance does not significantly improve between the 200ms and 400ms conditions (p > .22). The only significant performance improvement between these durations is for contour preserving stimuli. As the mean uncorrected detection rate is approximately 65% for adjacent contour violating stimuli at this rate, it is safe to assume that this is not a consequence of ceiling effects. Interestingly, detection of contour preserving changes does not significantly improve between either the 50ms and 100ms conditions or the 100ms and 200ms conditions.

On the other hand, the results do not support our first hypothesis. Because performance was at floor in the 50ms condition, we cannot characterize the relative contributions of adjacent and non-adjacent contour information to similarity judgments at this rate. Contrary to our prediction, the pattern of performance in the 100ms, 200ms, and 400ms tone duration conditions was identical: adjacent contour violating changes were detected best, followed by non-adjacent contour violations, with contour preserving changes least frequently detected.

We predicted a lack of adjacent contour advantage in the short duration conditions in part because of an expected streaming effect. At fast speeds, we expected that pitch proximity-based streaming effects would disrupt processing of temporal order information. As a result, changes that either altered within-stream order or added or removed an element from one of the streams would be more salient than temporal order changes that preserved within-stream order and kept the number of elements per stream constant. Assuming that these streaming changes were evenly distributed between adjacent- and non-adjacent-contour violating stimuli, the difference in sensitivity to changes in these stimulus classes should have been reduced. Unfortunately, the method we used to generate and select tone sequences for inclusion in the experiment did not result in stimuli that were readily quantifiable in terms of propensity to induce streaming effects. A similar experiment in which stimuli were designed with stream segregation in mind might help to clarify the extent to which this phenomenon interferes with contour processing.

CONCLUSIONS

Performance in the 50ms/tone condition strongly suggests that frequency processing at this rate is not representative of music processing. Our results do not exclude the possibility that repeated exposure to a stimulus at this presentation rate may allow for better encoding and therefore better change detection. However, this is not equivalent to the near-immediate extraction of contour information that occurs at slower rates. Continuing improvements in change detection between the 100ms/tone and 200ms/tone presentation rates suggest that results may generalize best if the interval between tone onsets is in an approximate range of 200 to 400 ms.

Our results also provide support for sensitivity to non-adjacent contour information in the general population. Our subjects had considerably less musical training than Quinn's, yet displayed the same general pattern of performance. A minor difference in results is present in that Quinn's subjects were more sensitive to changes to non-adjacent contour information between notes three notes apart (e.g. positions 1 and 4). Our subjects showed no significant differences between the two types of non-adjacent contour alterations, and had a non-significant tendency for superior performance in the two notes apart condition (e.g. positions 1 and 3).

While the presentation rate we used is relatively characteristic of natural music listening, the stimuli were not reminiscent of any natural musical genre. While the unfamiliar tonality of the Bohlen-Pierce scale limited interference from any expectancies about interval relations, it also limits the generalizability of our results. A carefully designed study utilizing stimuli derived from the Western tonal system would help to clarify the true relative importance of adjacent and non-adjacent contour information to melodic processing in the general population.

REFERENCES

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-36.

Bregman, A.S. (1990). *Auditory scene analysis: the perceptual organization of sound.* Cambridge, MA: The MIT Press.

Boltz, M., Marshburn, E., Jones, M. R., & Johnson, W. W. (1985). Serial-pattern structure and temporal-order recognition. *Perception & psychophysics*, *37*(3), 209–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4022750

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & psychophysics*, *50*(4), 305–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1758762

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *The Journal of the Acoustical Society of America*, *49*(2), Suppl 2:524+. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5541747

London, J. (2002). Cognitive Constraints on Metric Systems: Some Observations and Hypotheses. *Music Perception*, *19*(4), 529–550. doi:10.1525/mp.2002.19.4.529

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming tnumbers into movies, *Spatial Vision 10*, 437-442.

Quinn, I. (1999). The Combinatorial Model of Pitch Contour. *Music Perception*, *16*(4), 439–456.

Schröger, E, Näätänen, R., & Paavilainen, P. (1992). Event-related potentials reveal how non-attended complex sound patterns are represented by the human brain. *Neuroscience letters*, *146*(2), 183–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1491786

Schröger, Erich. (1994). An event-related potential study of sensory representations of unfamiliar tonal patterns. *Psychophys*, *31*, 175–181.

Sussman, E., Winkler, I., Ritter, W., Alho, K., & Näätänen, R. (1999). Temporal integration of auditory stimulus deviance as reflected by the mismatch negativity. *Neuroscience letters*, *264*(1-3), 161–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10320039

Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior Formation of Cortical Memory Traces for Melodic Patterns in Musicians. *Learning & memory (Cold Spring Harbor, N.Y.)*, *8*, 295–300. doi:10.1101/lm.39501.asked