# Melody Recognition after Short Delays: Effects of Contour Complexity and Key-Distance

Thomas W. Reiner
*Troy University*

## Abstract

*This study investigated melody recognition with a modification of the standard-comparison paradigm. Subjects listened to an original melody, were exposed to a silent retention interval, and then were presented with a target and a distractor. There were two types of trials. On target-same trials, listeners heard the target (the original melody) played in the same key exactly as it was previously heard and a distractor (a novel melody) played in a key either a major second or a perfect fourth from the original. On target-different trials, targets were heard in keys either a major second or a perfect fourth from the original, while distractors were played in the same key as the original. The contours of targets and distractors were also examined. Retention intervals varied from 0.5 s to 15 s. The results indicate that contour complexity and key-distance interact in the recognition of short melodies.*

An important goal in the field of music cognition is to understand how people perceive and recognize melodies. A melody is a succession of pitches in time. Melodies are encoded in an interval code and a contour code (Dowling, 1978; Edworthy, 1985). The interval code is related to the distance between any two adjacent tones in a melody along a logarithmic frequency scale. Intervals are the basis of musical scales and harmony. Contour, on the other hand, involves the pattern of up and down directions in a particular melody (Dowling, 1978; Edworthy, 1985; Massaro, Kallman, & Kelly, 1980). Together, interval and contour define a melody. Behavioral research has demonstrated that interval and contour are implicated in melody recognition (Cuddy & Lyons, 1981; Dowling & Bartlett, 1981; Dowling & Fujitani, 1971; Idson & Massaro, 1978; Massaro et al., 1980). Clearly, if we are to understand how melodies are perceived, we need to understand how interval and contour information are processed.

Melodies can be transposed to different keys. Transposition occurs when a melody is shifted upward or downward in pitch while the intervals between pitches remain constant (Dowling & Bartlett, 1981; van Egmond & Povel, 1996). A transposed melody sounds similar to a nontransposed version except that it will be higher or lower in pitch. Transpositions can either be exact or inexact. An exact transposition retains the interval distance relationship between tones, whereas the interval structure in inexact transpositions is changed. When we hear familiar melodies, they are usually exact transpositions to arbitrary pitch levels (Dowling, Kwak, & Andrews, 1995). We recognize exact transpositions as the same melody, regardless of the key being played. This makes transpositions ideal for studying melody recognition.

The distance of a transposition from an original melody can be measured in key-distance as well as pitch-distance. Key-distance is a measure of the extent that two keys share identical pitches. Key-distance is better understood if the circle of fifths is discussed. The circle of fifths describes a relationship between keys that shows that “close” (or near) keys have more notes in common with each other than “far” keys. If one takes a melody in C major for instance, the closest related key is G major. G major begins on the fifth note of the key of C and has only note that differs from C major, that of F#. The next key in the circle of fifths is D major, which begins on the fifth note of the G major scale. D major shares every note in common with G major except for C#. This progression continues as the fifth note of the current scale becomes the first note of the next scale through the entire cycle of fifths.

*Correspondence concerning this article should be addressed to Thomas W. Reiner, Department of Psychology, Troy University, 136 Catoma Street, Montgomery, AL, 36104. Email: treiner@troy.edu.*

Several studies have shown that identifying transposed melodies involves a key-distance effect where transpositions to near keys are more easily recognized than transpositions to far keys (Cuddy, Cohen, & Mewhort, 1981; Cuddy, Cohen, & Miller, 1979; Takeuchi & Hulse, 1992; Trainor & Trehub, 1993). Distracters or lures that are in "near" keys can be confused with target melodies, while distracters or lures in "far" keys are less confused with targets. This perceptual key-distance effect has been documented in several studies with adults (Bartlett & Dowling, 1980; Cuddy & Cohen, 1976; Cuddy et al., 1981; Takeuchi & Hulse, 1992, van Egmond & Povel, 1994) as well as kindergarteners and grade-school children (Bartlett & Dowling, 1980) and even infants (Trainor & Trehub, 1993). However, some researchers have failed to find a key-distance effect (van Egmond & Povel, 1994) and others maintain that key-distance effects may not be very robust or are due to the general context in which a melody is presented (Takeuchi & Hulse, 1992).

Pitch-distance is a measure of how far a given note is from another in half-steps. If one starts with C, the note D is two half-steps away. The note G is seven half steps away. Therefore, D is closer in pitch-distance than G. However, the key of G major is closer to C major than D major as specified by the circle of fifths. Pitch distance has been the focus of only a few studies (van Egmond & Povel, 1994; Takeuchi, 1992, Reiner, 2011). It has a larger effect on similarity judgments of melodies than key-distance (van Egmond & Povel, 1996). It is important to note that key distance and pitch distance are, to some extent, related and cannot be varied entirely independently (van Egmond, Povel, & Maris, 1996).

As mentioned earlier, contour is the other code that defines a melody. Contour can influence how listeners perceive a melody. Melodies that share similar contours and that are played in near keys from each other can easily be confused (Bartlett & Dowling, 1980). Melodies with few contour changes have simple contours and those that have more contour changes have complex contours. Contour complexity can be thought of as the number of contour changes in a melody. The findings on contour complexity are mixed. Jansen and Povel (2000) had listeners rate how good 6-tone sequences were as melodies. Listeners rated the musical goodness of melodies with simple contours higher than those with complex melodies. Additionally, some research indicates that people are able to recognize melodies with simple contours better than others with more complex contours (Cuddy & Lyons, 1981; Cuddy et al., 1981), while other studies suggest that the number of directional contour changes does not influence recognition (Croonen, 1994; Croonen and Kop, 1989). One way to examine this issue more clearly would be to test melody recognition with different delays and to vary the contour of targets and distracters on each trial. This would allow us to discern the extent that contour complexity contributes to melody recognition.

Melody recognition has been studied in many different ways. Some have tested whether people can identify distorted melodies (White, 1960), whether they can transpose six-note melodies to new keys using sine-wave oscillators (Attneave & Olson, 1971), or whether they can recognize familiar melodies that have been octave scrambled (Idson & Massaro, 1978; Massaro et al., 1980; Dowling, 1978). The most common method of testing melody recognition is the standard-comparison paradigm. (Bartlett & Dowling, 1980; Cuddy & Cohen, 1976; Dowling & Bartlett, 1981; Dowling & Fujitani, 1971; Dowling, 1978; Takeuchi & Hulse, 1992). This methodology involves playing a standard melody followed by a comparison melody and then having subjects indicate if the comparison was the same as or different from the standard (Dowling, 1978; Dowling & Fujitani, 1971). Sometimes pairs of comparisons melodies are used that consist of an exact transposition of the melody or an inexact transposition with at least one "wrong" note, and a novel melody, or "lure" (Bartlett & Dowling, 1980; Dewitt & Crowder, 1986; Dowling & Bartlett, 1981). The lure is incorporated in this design to ensure that the melody judgment is not a trivial task. In some cases, the contour of lures is varied so that they are either similar or different from the standard (Dowling, 1991; Watkins, 1985).

A modification of the standard-comparison paradigm is the similarity-comparison paradigm (van Egmond and Povel, 1996; van Egmond et al., 1996). In this case, two combinations of sequences are played. The first combination starts by playing a V-I cadence, then the standard melody, and then a transposition of the standard. The second combination is played after a short delay. This time the sequences are the cadence, the standard, and a second transposition of the standard. Listeners then choose which combination of melodies is more similar. Since lures are not introduced, as in the standard-comparison paradigm, subjects can make direct estimates of the similarity between transpositions (van Egmond et al., 1996).

Another approach to testing melody recognition involves using either a nontransposed melody or an exact transposition (inexact transpositions are not included), along with a novel melody, as the comparison melodies on each trial (Radvansky & Potter, 2000; Radvansky, Fleming, & Simmons, 1995; Reiner, 2011; Wolpert, 1990). Basically, each trial starts by presenting an original melody (referred to as the standard in other methodologies) in a given key and then playing two comparison melodies after a retention interval has elapsed. Radvansky and Potter (2000) tested four different comparison conditions. The first condition presents the original melody (the target) in the same key it was previously heard and a novel melody (the distracter) played in the same key as the original melody – all three melodies are played in the same key. This is referred to as a match-same trial (i.e., the keys of the target and original matched, and the keys of the target and distracter were the same). The second condition presents the target in the key of the original and a distractor in a different key from the target. This is called a match-different trial. The third condition presents the target in a different key than the original and the distracter in the same key as the target; this is called a mismatch-same trial. The fourth condition presents the target in a different key than the original and the distracter in the same key as the original; this is called a mismatch-different trial. After hearing the two melodies, a two-alternative forced choice (2AFC) was given to listeners that required them to choose which melody (the target or distracter) was the same as the original melody.

Reiner (2011) made four modifications to Radvansky and Potter's methodology. First, only Radvansky and Potter's (2000) second and fourth conditions were tested in which the keys of the target and distracter were always different from each other. Testing only these two conditions provides a focused means of examining the relationship of pitch and contour with the target and distracter. Second, the terminology was changed to only describe the key relationship of the target and original. Match-different trials were renamed as target-same (i.e., the key of the target was the same as that of the original), and mismatch-different trials were renamed target-different (i.e., the keys of the target and original were different). Third, the beginning of each trial included a V-I cadence to induce key (van Egmond & Povel, 1996). Fourth, since participants could use a guessing strategy with a 2AFC methodology, participants were asked to rate their level of confidence in their choice at the end of each trial. Even though the target is present on every trial, identifying the correct melody is a nontrivial task. The strategies used to identify unaltered versions of melodies are different from those used to identify altered ones (Schulkind, Posner, & Rubin, 2003). Additionally, the contours of the target and distracter can be manipulated to examine the effect of contour complexity and the pitch interval of transposition can be varied to assess the effects of pitch-distance and key-distance.

The purpose of the current study was to examine melody recognition after delays of 0.5 s to 15 s. A previous study that used identical methodology tested melody recognition with silent retention intervals of 3, 6, and 9 s (Reiner, 2011). The current study tested melody recognition with 0.5, 12, and 15 s silent retention intervals. The data from these additional retention intervals were combined with the earlier data and presented as one study. It was hypothesized that the accuracy rate for the 0.5 retention interval would be at least as high as the 3 s group. It was also hypothesized that accuracy would continue to decline after 9 s with the 12 s and 15 s retention intervals. The effects of pitch and contour were also assessed to see how these factors influence recognition and whether they interact. Also of interest was whether effects of contour complexity will be detected with silent retention intervals. Previous research with filled retention intervals (Cuddy et al., 1981) found that contour complexity influences recognition, while other studies (Croonen, 1994; Croonen & Kop, 1989) that used silent retention intervals did not find evidence of these effects.

## Method

### Participants

The sample included 180 undergraduate psychology students (49 men and 131 women), all with normal hearing, from universities and colleges in Alabama and Florida (ages 18-61 years, $M$ = 26.62, $SD$ = 9.6). Students participated voluntarily and received partial course credit in exchange for their participation. Based on their responses to a music training questionnaire, participants reported an average of 2.4 years ($SD$ = 3.7) of prior musical training which ended 4.6 years ($SD$ = 7.4) prior to the experiment. See the Appendix for the *Music Training Questionnaire*.

### Apparatus

The melodies were played on computers through 16-bit sound cards. Participants listened to the audio stimuli via Sony MDR-V150 headphones. The experiment was presented to participants by means of a Visual Basic application installed on each computer that guided them through the experiment and recorded their responses.

### Melodies

Sixty-four melodies played in the timbre of an acoustic piano were used as the auditory stimuli. Melodies were tonal and did not include any altered tones (accidentals) from the scales they were derived. Each melody was comprised of six tones of equal duration followed by a seventh tone that was twice the duration of each of the first six notes.  Each melody had a tempo of 150 beats per minute and was 6.4 s in duration.  The first six notes were each played for 800 ms and the seventh note was played for 1600 ms.  Examples of the melodies used in the study are shown in Figures 1 and 2.

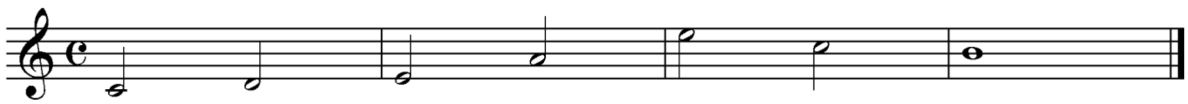

*Figure 1*. Example of a melody with a simple contour.

*Figure 2*. Example of a melody with a complex contour.

Each melody was equated for loudness, pitch range, and rhythm.  Stimuli were entered into a standard music notation program and digitally recorded at a constant dynamic level as audio files on a PC with a Pentium 4 processor.  Audio stimuli were recorded at a 44.1 kHz sampling rate with a 16-bit amplitude resolution.

Thirty-two melodies were designated as original/target melodies and 32 were distracter melodies.  Target melodies were identical to originals and were either heard in the same key as when they were initially played or they were transposed to a key that was a major second or a perfect fourth from the original key.  The transpositions used in the current study were "exact" so the intervals from the original melody were preserved.  The directions of the transpositions were varied so that half were upward and half were downward.  Distracter melodies were novel melodies that were different from the original melodies.  Distracters were either played in the same key as the original melody (with a starting tone of C) or they were played in keys a major second or a perfect fourth from key of the original melody.

A primary variable of interest in the study was "pitch-matching condition."  This condition corresponded to the particular sequence of keys the target and distracter were played in on a given trial.  There were two scenarios as mentioned earlier.  In the first scenario, the target was played in the same key as the original melody and the distracter was played in a different key.  This was referred to as a target-same trial.  In the second scenario, the target was transposed to a different key than the original melody was played in and the distracter was played in same key as the original melody.  This was referred to as a target-different trial.  Pitch matching condition was counterbalanced across trials so that half of the time listeners were given a target-same trial and half of the time they were given a target-different trial.  The specific keys that correspond to each melody were also counterbalanced across trials.  Refer to Table 1 for the order of the keys in each type of trial.

Table 1
*Order of Keys in Each Type of Trial*

| | Key of Each Melody | | |
|---|---|---|---|
| Trial | Original | Target | Distracter |
| Target-Same | Key 1 | Key 1 | Key 2 |
| Target Different | Key 1 | Key 2 | Key 1 |

*Note*. Target melodies were identical to original melodies except that they were transposed to different keys in target-different trials. Distracter melodies were novel melodies that were different from original melodies.

## Procedure

Each experimental session was conducted the same way.  The experimenter randomly assigned each participant to a given retention interval.  Participants were given a one-page informational document to read that explained the purpose of the study. After they were done reading the document, participants provided their informed consent to be in the study and then completed a short music training questionnaire. After participants completed the questionnaire, the experimenter described how to work through the experiment on the computer. The experimenter explained that for every trial they were going to listen to a melody and after a short delay would they hear two melodies: One which would be the same as the first one they heard, and a new melody. They would then be asked to identify which melody was the same as the first one they heard. It was also explained that the melody might be played at a higher or lower pitch than the first time they heard it.

Listeners put on audio headphones and worked through the experiment on computer workstations. The experiment consisted of 32 trials. Each trial started with a dominant to tonic chord cadence (V-I) in C major to induce the key. The G major chord was played for 1 s followed by the C major chord played for 3 s. After the cadence was played, a 6.4 s melody designated as the "original" was played. This was followed by showing a picture of the Grand Canal of Venice on the computer screen. The duration of the picture on the screen corresponded to the particular retention interval which each participant was randomly assigned (0.5 s, 3 s, 6 s, 9 s, 12 s, or 15 s).

After the picture was shown, participants heard two melodies successively separated by a 500 ms inter-stimulus interval. One of the two melodies was identical to the original melodies; it was designated as the "target" melody. The other melody was a novel melody and was referred to the "distracter." The presentation order of target and distracter melodies was counterbalanced across trials. Participants were instructed to choose the melody that matched the one they heard at the beginning of the trial. The procedure was a 2AFC so participants had to make a choice. After listeners made their choice, they were instructed to rate their level of confidence in their choice on a 5-point Likert scale ranging from "Not at all Confident" to "Extremely Confident." Higher numeric ratings corresponded to greater confidence.

## Musical Keys

The key of original melodies played at the beginning of each trial was always C major. Melodies that initially ascended in pitch started on middle C (261.63 Hz). Melodies that initially descended in pitch started on the "C" an octave above middle C (523.25 Hz). Targets and distracters that were played a major second above the key of original melodies (C major) were in the key of D major; melodies heard a major second below originals were in the key of Bb major; melodies heard a perfect fourth above originals were in the key of F major; and melodies heard a perfect fourth below originals were in the key of G major.

## Melodic Contour

The melodic contour of each melody was manipulated on each trial to examine whether simple or complex contour influenced melody recognition. Simple contour melodies either had one or two reversals, whereas complex contour melodies had three or four reversals. The contours of the target and distracter in each trial were combined into a contour pair. There were four possible target-distracter contour pairings: Simple-simple, simple-complex, complex-simple, and complex-complex. The pairing of contours was of interest since prior research suggests that a novel melody that has a contour similar to an original melody and that is also in a near-key from the original melody may be confused with the original melody (Bartlett & Dowling, 1980).

# Results

A between-within subjects analysis of covariance (ANCOVA) was conducted on the melody recognition data. Retention interval was a between-subjects factor, while pitch matching condition, pitch interval, and contour complexity were within-subjects factors. The number of years of prior musical training that each participant reported was the covariate. The dependent variable was the number correct on a series of 2AFC tests. All pairwise comparisons included Bonferroni correction. A significance level of $p < .05$ was used to assess all the statistical tests reported here. The mean proportion correct (where chance is 0.50), standard errors, and sensitivity measures (d' values) for each retention interval, pitch matching condition, pitch interval, and contour complexity combination are shown in Table 2.

Table 2
*Mean Proportion Correct, Standard Errors, and d' Values for Each Retention Interval, Pitch Matching Condition, Pitch Interval, and Contour Pair*

| | $n^a$ | *M* | *SE* | *d'* |
|---|---|---|---|---|
| Retention Interval | | | | |
| 0.5 seconds | 30 | .81 | .02 | 1.24 |
| 3.0 seconds | 30 | .78 | .02 | 1.09 |
| 6.0 seconds | 30 | .79 | .02 | 1.14 |
| 9.0 seconds | 30 | .79 | .02 | 1.14 |
| 12.0 seconds | 30 | .76 | .02 | 1.00 |
| 15.0 seconds | 30 | .71 | .02 | 0.78 |
| Pitch-Matching Condition | | | | |
| Target-Same | 180 | .84 | .01 | 1.40 |
| Target-Different | 180 | .70 | .01 | 0.73 |
| Pitch Interval | | | | |
| Major 2nd | 180 | .80 | .01 | 1.18 |
| Perfect 4th | 180 | .75 | .01 | 0.94 |
| Contour Pair | | | | |
| Simple-Simple | 180 | .79 | .01 | 1.14 |
| Simple-Complex | 180 | .77 | .01 | 1.04 |
| Complex-Simple | 180 | .83 | .01 | 1.34 |
| Complex-Complex | 180 | .71 | .01 | 0.78 |

[a] *n* refers to the number of people in each condition.

The between-subjects effect of retention interval was significant [$F(5, 173) = 2.54$, $p = .03$, $\eta_p^2 = .07$]. Listeners identified target melodies more often following a 0.5 retention interval ($M = .81$) than after a 15 s retention interval ($M = .71$). The covariate was also significant, $F(1, 173) = 10.73$, $p = .001$, $\eta_p^2 = .06$. Accuracy was associated with more years of prior musical training on target-same trials [$r(180) = .21$, $p = .006$] and on target-different trials [$r(180) = .26$, $p = .003$].

The within-subjects effect of pitch matching condition was significant [$F(1, 173) = 89.11$, $p < .001$, $\eta_p^2 = .34$]. Listeners identified target melodies more when they were heard in the original key ($M = .84$) than a different key ($M = .70$). The within-subjects effect of pitch interval was also significant [$F(1, 173) = 8.95$, $p = .003$, $\eta_p^2 = .05$]. Listeners demonstrated greater recognition when targets and distracters were played a major second from the original key ($M = .80$) than a perfect fourth ($M = .75$).

The within-subjects effect of contour complexity was significant [$F(3, 519) = 14.16$, $p < .001$, $\eta_p^2 = .08$]. Listeners recognized targets more when the contour combination was simple-simple, simple-complex, or complex-simple than when it was complex-complex ($M = .79$, .77, .83, and .71, respectively). Target melodies were also more recognizable when the contour combination was complex-simple ($M = .83$) than when it was simple-complex ($M = .77$).

All of the two-way interactions involving pitch matching condition were significant. See Table 3 for the mean proportion correct, standard errors, and sensitivity values (d' values) for each retention interval, pitch interval, and contour pair by pitch matching condition. The two-way interaction between pitch matching condition and retention interval was significant, $F(5, 173) = 3.04$, $p = .012$, $\eta_p^2 = .08$. Recognition was greater for target-same trials than for target-different trials across all retention intervals. For target-same trials, people were better able to recognize the target with 0.5, 3, 6, and 9 s retention interval ($M = .91$, .88, .86, and .85, respectively), than with a 15 s retention interval ($M = .74$). Listeners were also more accurate after a 0.5 s retention interval ($M = .91$) than after 12 s ($M = .82$) although this difference was only marginally significant. The accuracy rates on target-different trials did not differ significantly across retention intervals.

Table 3

*Mean Proportion Correct, Standard Errors, and d' Values for Each Retention Interval, Pitch Interval, and Contour Pair by Pitch Matching Condition*

| | Pitch Matching Condition | | | | | |
|---|---|---|---|---|---|---|
| | Target-Same | | | Target-Different | | |
| | *M* | *SE* | *d'* | *M* | *SE* | *d'* |
| Retention Interval | | | | | | |
| 0.5 seconds | .91 | .02 | 1.89 | .70 | .03 | 0.73 |
| 3.0 seconds | .88 | .02 | 1.24 | .68 | .03 | 0.66 |
| 6.0 seconds | .96 | .02 | 1.52 | .72 | .03 | 0.82 |
| 9.0 seconds | .95 | .02 | 1.47 | .72 | .03 | 0.82 |
| 12.0 seconds | .82 | .02 | 1.30 | .71 | .03 | 0.73 |
| 15.0 seconds | .74 | .02 | 0.90 | .67 | .03 | 0.66 |
| Pitch Interval | | | | | | |
| Major 2nd | .85 | .01 | 1.47 | .74 | .01 | 0.90 |
| Perfect 4th | .84 | .01 | 1.40 | .66 | .02 | 0.58 |
| Contour Pair | | | | | | |
| Simple-Simple | .90 | .01 | 1.80 | .68 | .02 | 0.66 |
| Simple-Complex | .87 | .01 | 1.59 | .66 | .02 | 0.58 |
| Complex-Simple | .85 | .01 | 1.47 | .80 | .02 | 1.18 |
| Complex-Complex | .75 | .02 | 0.94 | .66 | .02 | 0.58 |

[a]Each pitch interval occurred eight times in each pitch matching condition.
[b]Each contour complexity combination occurred four times in each pitch matching condition.

Paired t-tests were performed to compare the accuracy rates of the target-same and target-different trials for every retention interval. The target-same and target-different accuracy rates were significantly different from each other for every retention interval except 15 s [$t(29) = 1.97$, $p = .06$]. This showed that the delay of 15 s reduced the effect of hearing the target in the same key as the first time it was heard to the point that there was no difference from when the target was heard in a different key. Figure 3 shows the mean proportion correct for each pitch matching condition by retention interval.

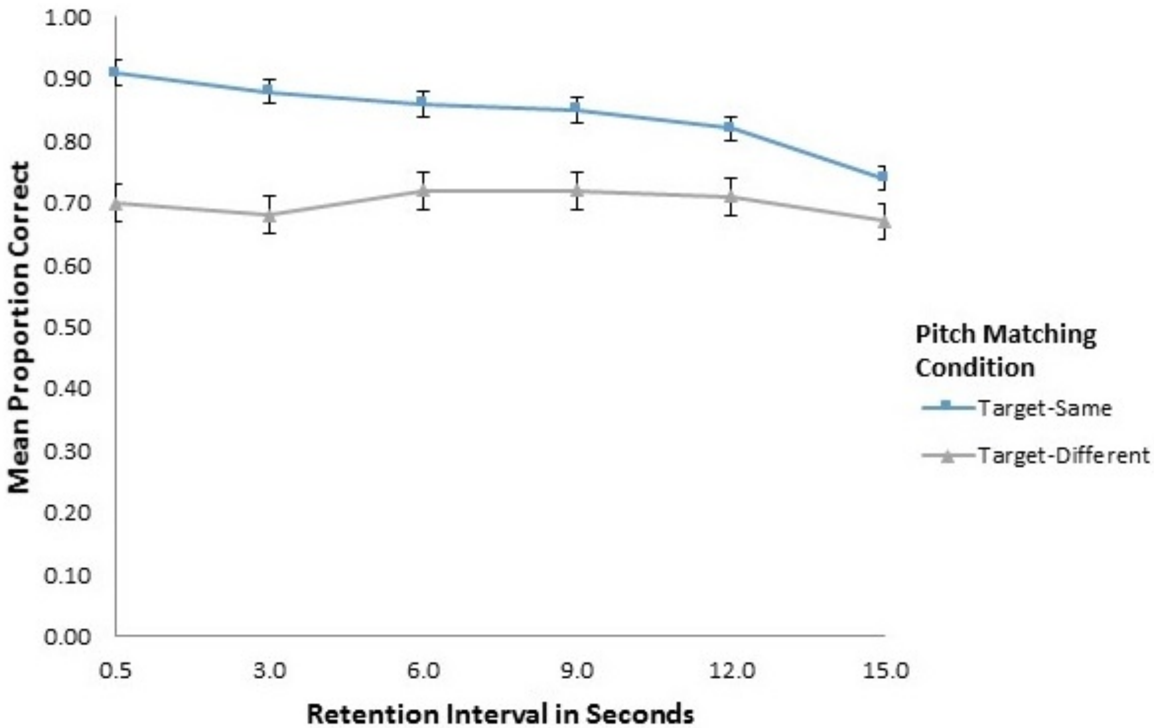


*Figure 3*. Mean proportion correct for each pitch matching condition by retention interval. Error bars represent standard errors.

The two-way interaction between pitch matching condition and pitch interval was significant, $F(1, 173) = 4.84$, $p = .03$, $\eta_p^2 = .03$. When the target was heard in the original key, there was no difference in recognition regardless if the distracter was heard a major second ($M = .85$) or a perfect fourth ($M = .84$) from the original key. A different situation resulted when the target was heard in a different key from the original key. In this case, recognition was higher when the target was heard a major second ($M = .74$) from the original key rather than a perfect fourth ($M = .66$) from the original key.

The two-way interaction between pitch matching condition and contour complexity was significant, $F(3, 519) = 12.39$, $p < .001$, $\eta_p^2 = .07$. Accuracy was higher when the target was heard in the original key and the contour pairs were simple-simple, simple-complex, and complex-simple ($M = .90$, .87, and .85, respectively) than when they were complex-complex ($M = .75$). Recognition was also higher for simple-simple contour pairs ($M = .90$) than with complex-simple ($M = .85$) contour pairs. When the target was heard in a different key than the original key, accuracy was higher with complex-simple contour pairs ($M = .80$) than simple-simple, simple-complex, or complex-complex contour pairs ($M = .68$, .66, and .66, respectively).

The two-way interaction between pitch interval and contour complexity was significant, $F(3, 519) = 16.14$, $p < .001$, $\eta_p^2 = .09$. The mean proportion correct, standard errors, and sensitivity measures (d' values) for each contour pair by pitch interval are shown in Table 4. When the target or distracter were heard a major second from the original key, recognition was greater with simple-simple, simple-complex, and complex-simple contour pairs ($M = .87$, .81, and .85, respectively) than with complex-complex pairs ($M = .66$). Accuracy was also higher with simple-simple pairs ($M = .87$) compared to simple-complex pairs ($M = .81$). When the target or distracter were heard a perfect fourth from the original key, recognition was greater with complex-simple pairs ($M = .80$) than with simple-simple or simple-complex pairs ($M = .71$ and .73, respectively).

Table 4
*Mean Proportion Correct, Standard Errors, and d' Values for Each Contour Pair by Pitch Interval*

| | Pitch Interval | | | | | |
|---|---|---|---|---|---|---|
| | Major Second | | | Perfect Fourth | | |
| | *M* | *SE* | *d'* | *M* | *SE* | *d'* |
| Contour Pair | | | | | | |
| Simple-Simple | .87 | .01 | 1.59 | .71 | .02 | 0.78 |
| Simple-Complex | .81 | .02 | 1.24 | .73 | .02 | 0.86 |
| Complex-Simple | .85 | .02 | 1.47 | .80 | .02 | 1.18 |
| Complex-Complex | .66 | .02 | 0.58 | .76 | .02 | 1.00 |

[a]Each pitch interval occurred sixteen times in an experimental session.
[b]Each contour pair occurred four times with each pitch interval.

The three-way interaction between pitch matching condition, pitch interval, and contour complexity was significant, $F(3, 519) = 7.05$, $p < .001$, $\eta_p^2 = .04$. See Table 5 for the mean proportion correct, standard errors, and sensitivity measures (d' values) of each pitch interval and contour pair by pitch matching condition. When the target was heard in the original key and the distracter was heard a major second from the original key, recognition was greater with simple-simple, simple-complex, and complex-simple pairs (*M* = .93, .87 and .88, respectively) compared to complex-complex pairs (*M* = .72). Accuracy was also higher with simple-simple contour pairs (*M* = .93) compared to simple-complex (*M* = .87) and complex-simple contour pairs (*M* = .88, although marginally significant). When the target was heard in the original key and the distracter was heard a perfect fourth from the original key, recognition was greater for simple-simple and simple-complex pairs (*M* = .87 and .87, respectively) compared to complex-complex pairs (*M* = .79).

Table 5
*Mean Proportion Correct, Standard Errors, and d' Values for Each Pitch Interval and Contour Pair by Pitch Matching Condition*

| | | Pitch Matching Condition | | | | | |
|---|---|---|---|---|---|---|---|
| | | Target-Same | | | Target-Different | | |
| Pitch Interval[a] | Contour Pair[b] | *M* | *SE* | *d'* | *M* | *SE* | *d'* |
| Major Second | | | | | | | |
| | Simple-Simple | .93 | .01 | 2.09 | .81 | .02 | 1.24 |
| | Simple-Complex | .87 | .02 | 1.59 | .75 | .02 | 0.94 |
| | Complex-Simple | .88 | .02 | 1.65 | .83 | .02 | 1.34 |
| | Complex-Complex | .72 | .02 | 0.82 | .59 | .03 | 0.32 |
| Perfect Fourth | | | | | | | |
| | Simple-Simple | .87 | .02 | 1.59 | .56 | .03 | 0.21 |
| | Simple-Complex | .87 | .02 | 1.59 | .58 | .03 | 0.28 |
| | Complex-Simple | .83 | .02 | 1.34 | .77 | .02 | 1.04 |
| | Complex-Complex | .79 | .02 | 1.14 | .73 | .02 | 0.86 |

[a]Each pitch interval occurred eight times within each pitch matching condition.
[b]Each contour pair occurred two times within each pitch interval and pitch matching condition.

When the distracter was heard in the original key and the target was heard a major second from the original key, recognition was greater for simple-simple, simple-complex, and complex-simple contour pairs (*M* = .81, .75, .83, respectively) than for complex-complex contour pairs (*M* = .59). Listeners were more accurate with complex-simple (*M* = .83) than simple-complex contour pairs (*M* = .75, although marginally significant). When the distracter was heard in the original key and the target was heard a perfect fourth from the original key, recognition was greater for complex-simple and complex-complex contour pairs (*M* = .77 and .73, respectively) compared to simple-simple and simple-complex (*M* = .56 and .58, respectively) contour pairs.

## Sensitivity Analysis

Signal detection theory is used to analyze data where there is a level of uncertainty. The typical paradigm asks participants to respond "yes" or "no" as to whether the target stimulus is present on a given trial. The

2AFC design used in this study differs from the typical single interval "yes-no" signal detection paradigm in that the target stimulus is present on every trial. As a result, only the sensitivity index (d') can be calculated.

The d' value for target-same trials was above threshold (d' = 1.40), as defined by d' > 1.0 (Macmillan & Creelman, 2005). However, the d' value for target-different trials did not reach threshold (d' = 0.73) indicating that listeners had more trouble discriminating the target from the distracter. This is similar to what other studies have reported that listeners are better able to recognize nontransposed melodies compared to those that have been transposed (Dowling & Bartlett, 1981; Dowling & Fujitani, 1971).

## Confidence Rating Analysis

Pearson correlations were calculated between participants' melody choice and their confidence rating on each trial to assess whether there was a relationship between confidence and accuracy. Correlations were also calculated for pitch interval, contour pair, and the number of years of musical training with confidence and accuracy. The correlations for each pitch matching condition are reported separately. Additionally, estimates of guessing based on listener confidence ratings are included for each pitch matching condition.

**Target-same condition.** Higher confidence ratings were associated with greater accuracy, $r(2880) = .30$, $p < .001$. The pitch of either the target or distracter was not associated with confidence [$r(2880) = -.03$, $p = .08$] or accuracy [$r(2880) = -.01$, $p = .44$]. Greater complexity of contour pairs (simple-simple, simple-complex, etc.) was associated with lower confidence ratings [$r(2880) = -.18$, $p < .001$] and less accuracy [$r(2880) = -.14$, $p < .001$]. More years of prior musical training were associated with greater confidence [$r(2880) = .19$, $p < .001$.] and accuracy [$r(2880) = .08$, $p < .001$].

Guessing was estimated by examining confidence ratings and accuracy. Responses of "Not at all confident" and "Somewhat not confident" were combined into a "low confidence" category. Responses of "Moderately confident" represented a "moderate confidence" category. Lastly, responses of "Somewhat confident" and "Extremely confident" were combined into a "high confidence" category. Of the total number of correct choices made on target-same trials, 73.1% of selections were rated with high confidence, 17.5% were rated with moderate confidence, and 9.1% were rated with low confidence. Listeners were not as confident when their choice was incorrect. Of the total number of incorrect choices made on target-same trials, 39.7% were rated with high confidence, 31.9% were rated with moderate confidence, and 28.4% were rated with low confidence.

**Target-different condition.** Higher confidence ratings were associated with greater accuracy, $r(2880) = .22$, $p < .001$. Hearing the melody played a major second from the original key than a perfect fourth resulted in greater confidence [$r(2880) = -.07$, $p < .001$] and accuracy [$r(2880) = -.09$, $p < .001$]. There was no relationship between contour pair and confidence [$r(2880) = -.02$, $p = .23$] or accuracy [$r(2880) = .02$, $p = .43$]. Lastly, more years of prior musical training were associated with greater confidence [$r(2880) = .19$, $p < .001$] and accuracy [$r(2880) = .08$, $p < .001$].

Of the total number of correct choices made on target-different trials, 63.5% were rated with high confidence, 22.4% were rated with moderate confidence, and 14.1% were rated with low confidence. The pattern of confidence ratings for incorrect choices on target-different trials was similar to what was seen for the target-same trials: 41.5% of choices were rated with high confidence, 32.5% were rated with moderate confidence, and 26.0% were rated with low confidence. Listeners were less confident in their choices on target-different trials, regardless if they chose correctly or not.

# Discussion

This study extended the previously published work by Reiner (2011). Nontransposed melodies were recognized more often than when they were transposed, although this effect dissipated somewhat over the course of 15 s. The addition of the 0.5 s delay to the previous data provided a baseline for comparison and established a ceiling effect for this particular experimental design. Recognition with a 0.5 s delay was higher than the 3 s delay which was hypothesized. However, there was no difference between the accuracy rates for the 0.5 or 3 s delays. Recognition with the added 12 and 15 s delays decreased beyond what was found with the 9 s delay which was also hypothesized, but again there were no differences in these accuracy rates. The findings also support the idea that contour complexity influences melody recognition with silent retention intervals.

Although people were less accurate on target-different trials, even with very short delays, it may not be that they were unable to recognize transposed melodies but that they were responding to the keys of original melodies. This could be due to how pitch and contour are processed, since pitch information is thought to be encoded automatically and contour through controlled processes (Dowling et al., 1995). If listeners were only attending to the pitch of the original melody, then they would respond the way they did in the current study and mistakenly choose the distracter more often on target-different trials. The decline of accuracy on target-same trials with delays greater than 9 s supports a decay of memory trace explanation. However, accuracy on target-different trials did not change much with increased delays. This suggests that the ability to recognize transposed melodies, although not as good as that of identifying nontransposed melodies, is relatively stable for at least 15 s. Given that a silent retention interval was used, the question arises why accuracy on target-same trials declined after 9 s.

The two major explanations for forgetting are interference and decay. Interference can be ruled out in the current study since a silent retention interval was used. Of course, this brings up the question of whether listeners could rehearse the melodies. While it is possible that listeners used a rehearsal strategy, there probably was not enough time for rehearsal to occur with retention intervals of 3 s or less. Rehearsal would have been more likely with longer delays. If listeners were implementing a rehearsal strategy, then their accuracy rates would have been expected to be near perfect and that was not the case. Novel melodies, such as those used in this study, may not lend themselves to being rehearsed. Listeners, instead, would need to rely on pitch or contour information to recognize the melodies.

A second explanation is that the memory trace decayed over time. There is some debate about whether memory traces of auditory stimuli can decay with short retention intervals, but most of the findings against decay involve verbal as opposed to nonverbal auditory memory (Berman, Jonides, & Lewis, 2009; Lewandowsky, Oberauer, & Brown, 2009). There are a few studies that document decay of nonverbal auditory memory. A mismatch negativity study involving standard and "deviant" tones (Sams, Hari, Riff, & Knuutila, 1993) showed that the auditory memory trace decays over a period of 10 s. McKeown and Mercer (2012) also found that memory for tones in a standard-comparison task with silent retention intervals was better for short (1 to 4 s) than long (8 to 32 s) intervals. These findings from these studies suggest that auditory memory can deteriorate without interference after about 10 s.

When listeners are asked to identify a melody they just heard, the pitches are fresh in short-term memory. Memory for nontransposed melodies is simply a matter of pitch recognition (Dowling & Fujitani, 1971), so contour information should not influence recognition with target-same trials because pitch recognition is all that is necessary to identify the melody. This may be the case with shorter delays and it fits will what was seen in the current study with retention intervals from 0.5 to 9 s. After a 9 s delay, however, accuracy on target-same trials began to decrease markedly which fits more with a decay of memory trace explanation than the explanation given by Croonen and Kop (1989) that a memory representation is not fully formed until around 15 s. The higher rate of recognition with a 15 s delay reported by Croonen and Kop may have been due to individual differences of the subjects being tested. There may have been more subjects with recent musical training in the 15 s delay group than in other groups.

On target-same trials, listeners demonstrated the highest rate of recognition when the distracter was played a major second from the original key and the contour pair was simple-simple. Recognition decreased as the contour combinations became increasingly complex with the complex-complex contour combination evoking the worst rate of recognition. Target-same trials in which the distracter was played a perfect fourth from the original key produced similar patterns of recognition except that the recognition rates of simple-simple and simple-complex pairs were the same. Recognition on complex-simple pairs was slightly lower and recognition on complex-complex pairs was, again, the lowest.

When it came to target-different trials with the target played a major second from the original key, the pattern of results was slightly different from what was seen with target-same trials. In this case, recognition was highest with simple-simple pairs, decreased with simple-complex pairs, increased with complex-simple pairs, and was the lowest with complex-complex pairs. This time, targets having more contour information than distracters made them more discernible. A different pattern showed up from the other conditions when targets were played a perfect fourth from the original key on target-different trials. In this case, melody pairs that had complex targets produced greater recognition than melody pairs with simple contours. This suggests a key-distance effect mediated by contour complexity. Listeners were better able to recognize targets transposed to near keys (but far pitch distance) but only when they had complex contours.

Van Egmond and Povel (1996) observed that pitch-distance has a larger effect on similarity judgments than key-distance. If this were the case in the current study then recognition would have been worse when targets were transposed to the interval of a perfect fourth than to a major second. Instead, it was the combination of contour along with key-distance that predicted recognition. Targets with simple contours were not recognized as well as those with complex contours. Dyson and Watkins (1984) showed that novel melodies that have contour reversals are more salient, or perceptually distinct, than melodies with nonreversals. So, complex contour targets may have been more recognizable because the changes in direction made them more perceptually distinct to listeners.

The fact that targets were recognized more often when they were transposed to the interval of a perfect fourth may be related to the sense of tonality established by the initial cadence and the key of the original melody. Previous work (Cuddy et al., 1981; Cuddy et al., 1979) showed that listeners can recognize melodic sequences transposed to the dominant (the fifth tone in a major scale) more readily than to the tritone (the interval of an augmented fourth). While transposition to the interval of the tritone was not tested in the current study, transposition down a perfect fourth was tested which is the same as transposition to the dominant (it is an inversion of a fifth above the tonic). Additionally, transposition up a perfect fourth is the same as transposition to the subdominant (the fourth tone in a major scale). The dominant and subdominant are the most common keys of harmonic progression in Western music. Progression to these two keys takes place in a large corpus of music. So, the effect of recognizing targets transposed to the interval of a perfect fourth (either up or down) may be the product of repeatedly being exposed to common harmonic progressions in Western music.

## Implications

These findings support the idea that pitch is processed automatically, while contour is processed by controlled processes. The greatest predictor of melody recognition was the pitch matching condition. Transpositions were more difficult for people to recognize across every retention time. When recognition simply involves identifying a previously heard melody, the pitch of a distracter melody does not interfere with recognition. Increasing the contour complexity of melody pairs decreases a listener's ability to discriminate a target from a distracter. When recognition involves identifying a transposition, it is helpful if there is more contour information present in the target compared to the distracter. Otherwise, pitch information may mislead someone to mistake a melody played in the key of an original melody as being the original melody.

These findings can also inform us about ways of learning and performing music. Melodies that are transposed to near keys, but are farther away in pitch distance, might be better discerned if they have a complex contour. This concept could be incorporated into the way that composers write music and in the way that students learn to play music that is in a contrapuntal style. A main melody or theme might be better recognized if it has a complex contour. However, this is not always the case since simple melodies can be very distinct and memorable. An example would be the well-known opening motif of Beethoven's Fifth Symphony. There may be no more highly recognizable theme than those four notes in Western music.

It should also be mentioned that the stimuli in this study cannot provide information about the temporal aspects of music such as tempo, rhythm, and duration. What this study does show is how melodic contour and pitch information interact when we control for the temporal aspects of melodies. In conclusion, this study showed that the memory trace for non-transposed melodies decays over time and the recognition of transposed melodies, while worse than that of non-transposed melodies, is relatively stable over 15 s. Understanding why the recognition of transposed melodies is stable during this interval of time warrants further research and will greater inform us about how people process melodic stimuli.

## References

Attneave, F., & Olson, R. K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *The American Journal of Psychology*, *84*, 147-166.

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 501-515.

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 317-333.

Croonen, W. L. M. (1994). Effects of length, tonal structure, and contour in the recognition of tone series. *Perception and Psychophysics*, *55*, 623-632.

Croonen, W. L. M., & Kop, P. F. M. (1989). Tonality, tonal scheme, and contour in delayed recognition of tone sequences. *Music Perception*, *7*, 49-68.

Cuddy, L. L., & Cohen, A. J. (1976). Recognition of transposed melodic sequences. *The Quarterly Journal of Experimental Psychology*, *28*, 255-270.

Cuddy, L.L., & Lyons, H. I. (1981). Musical pattern recognition: A comparison of listening to and studying tonal structures and tonal ambiguities. *Psychomusicology*, *1*,15-33.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 869-883.

Cuddy, L. L., Cohen, A. J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology*, *33*, 148-157.
Dewitt, L. A., & Crowder, R. G. (1986). Recognition of novel melodies after brief delays. *Music Perception*, 259-274.
Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341-354.
Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, *50*, 305-313.
Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology: A Journal of Research in Music Cognition*, *1*, 30-49.
Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, *49*, 524-531.
Dowling, W. J., Kwak, S., & Andrews, M. W. (1995). The time course of recognition of novel melodies. *Perception & Psychophysics*, *57*, 136-149.
Dyson, M. C., & Watkins, A. J. (1984). A figural approach to the role of melodic contour in melody recognition. *Perception & Psychophysics*, *35*, 477-488.
Edworthy, J. (1985). Interval and contour in melody processing. *Music Perception*, 375-388.
Idson, W. L., & Massaro, D. W. (1978). A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, *24*, 551-565.
Jansen, E. L., & Povel, D. J. (2000). The role of implied harmony in the perception of brief tone sequences. In *Proceedings of the Sixth International Conference on Music Perception and Cognition* (pp. 927-934).
Lewandowsky, S., Oberauer, K., & Brown, G. D. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences*, *13*, 120-126.
Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*. (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
Massaro, D. W., Kallman, H. J., & Kelly, J. L. (1980). The role of tone height, melodic contour, and tone chroma in melody recognition. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 77.
McKeown, D., & Mercer, T. (2012). Short-term forgetting without interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1057-1068.
Radvansky, G. A., & Potter, J. K. (2000). Source cuing: Memory for melodies. *Memory & Cognition*, *28*, 693-699.
Radvansky, G. A., Fleming, K. J., & Simmons, J. A. (1995). Timbre reliance in nonmusicians' and musicians' memory for melodies. *Music Perception*, *13*, 127-140.
Reiner, T. W. (2011, September). Pitch-distance and contour complexity in the recognition of short melodies. *Journal of Scientific Psychology*, 27-36.
Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence. *Journal of Cognitive Neuroscience, 5*, 363-370.
Schulkind, M. D., Posner, R. J., & Rubin, D. C. (2003). Musical features that facilitate melody identification: How do you know it's "your" song when they finally play it? *Music Perception*, *21*, 217-249.
Takeuchi, A. H., & Hulse, S. H. (1992). Key-distance effects in melody recognition reexamined. *Music Perception*, *10*, 1-23.
Trainor, L. J., & Trehub, S. E. (1993). Musical context effects in infants and adults: Key distance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 615-626.
van Egmond, R., & Povel, D. J. (1994). Similarity judgments on transposed melodies as a function of overlap and key-distance. In *Proceedings of the 3rd International Conference on Music Perception and Cognition, Liège* (pp. 219-220).
van Egmond, R., & Povel, D. (1996). Perceived similarity of exact and inexact transpositions. *Acta Psychologica*, *92*, 283-295.
van Egmond, R., Povel, D. J., & Maris, E. (1996). The influence of height and key on the perceptual similarity of transposed melodies. *Perception & Psychophysics*, *58*, 1252-1259.
Watkins, A. J. (1985). Scale, key, and contour in the discrimination of tuned and mistuned approximations to melody. *Perception & Psychophysics*, *37*, 275-285.
White, B. W. (1960). Recognition of distorted melodies. *The American Journal of Psychology*, *73*, 100-107.
Wolpert, R. S. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. nonmusicians. *Music Perception*, *8*, 95-105.