Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Memory for Sequence Order in Songs.

Craig P. Speelman,*[1] Susan Sibma*, Simon MacLachlan*

*School of Psychology and Social Science., Edith Cowan University, Australia
[1]c.speelman@ecu.edu.au

ABSTRACT

Previous research on memory for music has typically measured RT and accuracy in tests of recall and recognition of songs. Little research, however, has focused on the ability of people to switch their attention between various parts of a song to answer questions about those parts. One hypothesis is that, because music unfolds in time, one's ability to consider different parts of a song might be influenced by where in the song someone begins their consideration, and also in which direction they are then asked to switch their attention, with the overriding bias being in a forwards direction. The current study tested this forward bias hypothesis. Fifty people were asked to identify whether the second excerpt (*target* line) of a pair of excerpts taken from a song came 'before' or 'after' the first excerpt (*probe* line) in the normal course of the song. Seven pairs of excerpts, three pairs falling before the target line, and four pairs occurring after the target line, were presented for each of 8 popular and 2 new songs. It was predicted that RTs for identifying the target lines occurring 'after' the probe line would be shorter than those coming 'before' the probe line. Results supported this hypothesis. The familiarity of a song did not affect this result. A companion experiment that compared performance on this task for musicians and non-musicians replicated these results, but indicated no effect of musical expertise. These results support the hypothesis that memory for songs is biased in a forward direction.

I. INTRODUCTION

Although a great deal of research has been conducted on the factors that affect memory for music, there has been little attention on what occurs during the time that elapses between hearing a song and giving a response that the song has or has not been recognised. Presumably, participants hum or sing to themselves as they try to recall or recognise the song, however, no research has focused on the ability of people to switch their attention between various parts of a song in order to answer questions about those parts. Specifically, when a song is recalled, is the experience akin to pushing play on a tape recorder, which can be considered a form of serial access to the musical memory, or do we possess more flexible mental facilities with parallel access to all parts of a song, akin to the random access memory of a computer?

The research that has been conducted on memory for musical stimuli has tended to focus on the relationship between memory for music and lyrics of songs, and also the ability of music to assist in memory for text. For example, Rubin (1977) asked participants to recall the text (in writing) of "The Star Spangled Banner" in one of three conditions; in silence, with the correct music playing, or with another melody ("Stars and Stripes Forever") playing. Recall of the text in the correct condition was significantly better than the silent condition, which was in turn better than the incorrect music condition. Rubin explained the difference as being due to the correct rhythm accompanying the correct music condition consequently facilitating recall. In the incorrect music condition, however, the melody was a distractor and interfered with recall of the text.

Wallace and Rubin (1991) showed that the repetition of a melody across verses leads to improved memory for the words of the verses compared to verses with varying melodies. Participants were presented with three verses, each sung to a different melody but retaining the same metre, rhythm, timing, notes and tonal centre of the original melody. Verbatim recall in this condition was compared with a spoken condition and a condition in which all verses were sung to the same melody. It was found that recall for the conditions containing three different melodies was not significantly different from recall in the spoken condition, and that verbatim recall for the same-melody condition over three verses was significantly better than in the other two conditions. Wallace and Rubin concluded that when a new melody is first heard, it acts as a distractor hindering lyrical recall. However, as the melody is repeated across verses and becomes more familiar to the listener, recall is facilitated by providing structural information regarding syllables, accents and stress patterns within a line of the verse. Together, these functions of melodic repetition aid lyrical recall and help the listener to differentiate one verse from another. It was also found that repetition of rhythmical information (the consistency and regularity of a musical beat) across verses somewhat overcame the distracting influence of the melodies, helping to facilitate recall, and approximating recall for the spoken condition. Wallace (1994) later explained these results by invoking a "chunking" mechanism as an aid to memory for song, suggesting that individual pieces of information are collected both perceptually and cognitively as larger, integrated "wholes" (Miller, 1956). For melody and verbal material (text), chunking increases the encoding capacity of STM and improves both the storage and retrieval processes of LTM (Dowling, 1973). Wallace further suggested that the chunking process for songs works in a series of stages, with rhythm grouped in word pairs or phrases, which are in turn grouped by the melody and number of stresses per line. Rhyming lines are also linked together, with verses echoing the same idea also grouped together. Imagery and meaning expressed through this configuration of word pairing, rhyme and word stressing are integrated in order to help facilitate memory for the text.

Longitudinal studies on the development of the ability to reproduce songs agree in the order of appearance of the various parts of songs during learning (Gérard & Auxiette, 1992). The words are learned first, then the rhythm, the melodic contour, the intonation, and finally the integration of the various components of the piece. The extent to which a song can be remembered well depends on the level of prosodic match between the text and music. The better the match between accented and unaccented beats, the better the recall. Together, the stages leading up to the integration implies a relative dissociation of the musical components. During learning, priority is given to either melody or text depending on the difficulty of the task and the participant's musical training. From their experiment, Gérard and Auxiette concluded that when processing two temporal sequences of text and melody at a time, one sequence drives the other and also marks points upon which the events of the other component can be hung.

Rubin (1977) and Serafine, Davidson, Crowder and Repp (1986) support this theory, noting that songs, when remembered, were thought to be processed as prosodic effects of non-semantic sound patterns. For example, if the melodic component of a song is changed, the words sound different. Moreover, when the words of a song are altered, the melody sounds different. For example, “The ABC Song”, “Twinkle, Twinkle, Little Star”, and “Baa Baa Black Sheep” are songs that possess virtually the same melody, with the first two songs being melodically identical while the third is a subtle variation. And yet many people are surprised when this is pointed out to them.

It has been suggested by Wallace (1994) that music can structure recall for words so that when words are forgotten, memory for the music provides cues enabling the words to be ‘picked up’ again. Tan and Speckman (2005) propose that memory for words may structure recall for the melody. However, the melody has to be sufficiently simple to be useable as a framework for recalling words. Further, the participant must be sufficiently musically expert to learn and recall words accurately.

Importantly, Bartlett and Snelus (1980) found the melodies of songs were better cues to recall the lyrics of a song than the title was. According to Bartlett and Snelus, the difference in cue effectiveness might have been a reflection of the separateness of the components of melody and text that comprises song. The separateness of the components of melody and text was examined by Serafine, Crowder and Repp (1984) and extended by Serafine et al. (1986). The findings from these studies revealed that while text and melody might be learned and remembered independently of each other, when learned together as a song, they appear to be processed differently, forming an integrated memory representation.

It is clear from the research reviewed so far that the surface structure, rhythm, and chunking mechanisms of song are important in facilitating recall. However, Wallace (1994) suggested that repetition of melody across verses in a song is also important. The repetition clearly distinguishes each verse from the others, provides both a structural frame regarding the number of lines, and also to cues in relation to the number of syllables, accents, and stress patterns in a line. Together, these functions of melodic repetition serve to facilitate the recall of the lyrics. Furthermore, recall should improve as the song unfolds, so that recall of lines towards the end of the song should be better than for lines in the beginning of the song in both popular and novel songs.

Previous research findings on memory for music (Rubin, 1977; Wallace, 1994; Wallace & Rubin, 1991), and in particular, Frances’ (1958/1988, pp.202) assertion that listeners to music experience an irreversible succession of unfolding sounds, suggest that people are caught in a forward trajectory while linking past to present sounds.

An interesting observation was made by Rubin (1977) of the participants in the correct music condition of his experiment. These participants, without exception, adopted the same recall strategy. They would write as quickly as possible until the music got ahead of them, at which point they would wait until the music came around to where they had stopped writing, pick up the writing once more and continue in another burst until the music got ahead of them again. This strategy suggests the need to continue moving forward through the song, with the recall of text being facilitated by the melody.

The research on memory for music provides some suggestion that memory representations for musical stimuli possess a directional feature. That is, access to elements of a musical memory representation may be facilitated or inhibited depending on the degree to which the type of access matches the normal direction of the music. For instance, if someone is presented with an excerpt (probe line) from a song and then asked to identify whether a second excerpt (target line) comes before or after the first excerpt, memory search would be serial in nature, such that the song would be replayed in a forward direction from the probe line until the target line is reached. If the target line comes before the probe line in the normal course of the song, the memory search would progress from the probe line in a forward direction, reach the end of the song and then return to the beginning until the target line is located.

The aim of the current study was to investigate the direction of memory for music in order to test the hypothesis that memory for music is biased in a forward direction. It was hypothesised that reaction times in identifying song excerpt positions would be quicker for excerpts occurring after the probe line in the normal course of the song, and longer for excerpts occurring before the probe line. Specifically, it was hypothesised that RTs will increase incrementally the further ‘forward’ the target line is from the probe line. If this is in fact the case, then the mean reaction times (RTs) to identify a target line that falls before the probe line, in the normal course of the song, would be greater than RTs for target lines occurring after the probe line.

II. METHOD

A. Participants

A convenience sample of 66 people took part in this study, comprising 52 students from Edith Cowan University and 14 members of the general public. The study involved two stages: a song selection stage and an experimental stage. The 16 participants who took part in the first stage ranged in age from 15 to 51 years, and consisted of eight females (mean age = 26.13 years) and eight males (mean age = 31.38 years). The remaining 50 participants, none of whom had participated in the first stage, completed the experimental stage. The age range for these participants was 15 to 59 years. There were 26 females (mean age = 29.54 years) and 24 males (mean age = 31.25 years). Participation in the study was voluntary, and each participant in the experiment was given two tickets in a raffle for a prize of $50.

B. Materials

A selection of 20 popular songs was compiled from Hit Songs lists that appeared in the print media and on the internet, and were heard on the radio. Two recent compositions not yet heard by the general public served as new songs. All songs used in the study came from compact disc recordings.

The methodological constraints of the experiment required that each of the 22 songs had the same structure of three 4-line verses, and a 3- or 4-line chorus played at the beginning and/or end of the song, and which was repeated in between verses. None of the lines of any verse was repeated in any other verse or in the chorus. Recordings of two additional songs (Mary had a little lamb, Blowin' in the wind) were used as short practice songs. Although the practice songs were shorter in length than the experimental songs, they followed the same structure of the popular and new songs.

In the experimental task, two excerpts of a song were always played on each trial. The probe line was played first, and the target line was played second. Participants were required to indicate whether the target line came before or after the probe line in the original song. In each song, the probe excerpt was always the second line of the second verse. The target excerpts were the 'same' line (i.e., the second line) of the earlier and later verses, the 'next' line (i.e., the third line) in each of the three verses, the last line of verse three, and the first line of verse one.

The rationale for choosing the particular probe and target lines related to the fact that the melody for a particular line in any verse is identical for the corresponding line in other verses. Each Line 2 (the 'same' line) that was tested had the same melody as the probe line, and each of the Line 3 excerpts (the 'next' line) also shared a common melody.

By keeping the line constant (either Line 2 or 3) it was possible to test for the effect of verse position. For the Line 2 targets, the melody should have been recognised as being the same as the probe line, so the decision was to choose which of the other two verses the lyrics belonged to. In the case of Line 3, however, it was expected that comparing Line 3 targets with the probe line would be confusing for the Verse 1 and 3 excerpts, as the melody would continue from the probe to the target as expected, but the lyrics would belong to different verses.

In his experiment on very long-term memory for prose, Rubin (1977) found a strong primacy effect and also a recency effect. The first line of Verse 1 (the first line of the entire song for 18 of the 22 songs) and the last line of Verse 3 of the songs were included in the current experiment to test for evidence of the same finding in song. In addition to an expected primacy effect however, the fact that the first line of songs often includes the words from the song title, was expected to further facilitate recall of this line making it almost too easy to identify. Therefore, the RTs for Line 1 were not included in the analysis of RT but were retained for accuracy analysis to give a more complete picture of accuracy for the songs. Furthermore, the inclusion of Line 1 in the current experiment was also necessary to maintain the balance of the number of targets in the *Before* and *After* conditions. In Rubin's (1977) experiments the finding of a recency effect for prose was not as strong as the primacy effect. In song, the last line of a verse is rarely (if ever) the last line of the complete song (the chorus is usually repeated at the end), and possibly diminishes any recency effect. Therefore the last line was included for both RT and accuracy analyses.

Within each song, the mean length of the excerpts occurring before the probe line, and the excerpts that fell after the probe line were roughly equivalent. The mean length of the *Before* excerpts was 6694ms, and the mean length of the *After* excerpts was 6682ms. The lyrics for one of the popular songs (Bad Moon Rising) are presented in Table 1, with the probe and target lines indicated.

Table 1. Probe and target lines for one of the popular songs (Bad Moon Rising), with correct responses indicated in parentheses.

Verse 1	
I see the bad moon rising	**'first' line, earlier verse (*Before*)**
I see trouble on the way	**'same' line, earlier verse (*Before*)**
I see earthquakes and lightning	**'next' line, earlier verse (*Before*)**
I see bad times today	
Chorus	
Don't go around tonight	
Well, it's bound to take your life	
There's a bad moon on the rise	
Verse 2	
I hear hurricanes a-blowing	
I know the end is coming soon	**Probe Line**
I fear rivers overflowing	**'next' line, same verse (*After*)**
I hear the voice of rage and ruin	
Chorus	
Verse 3	
Hope you got your things together	
Hope you are quite prepared to die	**'same' line, later verse (*After*)**
Looks like we're in for nasty weather	**'next' line, later verse (*After*)**
One eye is taken for an eye	**'last' line, later verse (*After*)**
Chorus	

C. Procedure

In order to test the suitability of the songs selected for the experiment, 16 participants were asked to indicate, by checking a box next to the name of each song, which songs they were familiar with in regard to both the lyrics and the melodies. The mean number of songs recognised per person was 13, and each song was selected at least three times. This result suggested that all songs were appropriate stimuli for this experiment in that most people knew most of the songs.

The experiment consisted of an orientation phase and a test phase. The orientation phase was designed to reacquaint participants with the popular songs, introduce them to the new songs, and orient them to the mode of presentation and responding in the test phase. Participants were provided with an instruction sheet and the list of songs. Participants were instructed to select eight songs from the list whose lyrics they knew well enough to sing along with. After making their choice from the song list, all eight popular songs, the two new songs, and the practice songs were played in their entirety to the participants via headphones. This phase lasted 30-40 minutes depending on the individual's choice of songs.

When the orientation phase was completed, the practice sessions commenced. In order to make clear the mode of responding in the experimental task, an analogous task was developed using the standard number counting sequence (i.e., 1, 2, 3, etc) instead of music stimuli. Participants initiated this task by pressing the space bar on the computer keyboard. The practice task involved four trials. On each trial, two numbers were presented on the computer screen. Participants were required to decide whether the second of the two numbers came before or after the first number in a normal counting sequence. A response was made by pressing the 'Z' key on the keyboard (marked as "B" for 'before') or the '/' key (marked as "A" for 'after'). Following completion of the counting task, excerpts from the practice songs were presented. Automatic accuracy feedback was provided to the participants by the computer after each trial in the number and practice song tasks. When it was clear that the participants understood the nature of the task and were able to respond appropriately, the test phase commenced.

Trials in both the practice phase and the test phase followed the same format as the counting task but with excerpts from songs replacing the numbers. Participants initiated a trial by pressing the space bar on the keyboard. This resulted in a message appearing on the screen that remained visible throughout the trial. This message instructed participants to press a response key as soon as possible after making their decision: the key marked "B" if they thought the second excerpt came before the first excerpt, or the key marked "A" if they thought it came after the first excerpt. The probe line was always played first, and was preceded by the spoken phrase "excerpt one", while the phrase "excerpt two" was presented before each of the target lines. This procedure ensured a consistent comparison relationship between excerpts throughout the experiment.

The computer recorded RT as the time that elapsed between the presentation of the target excerpt and the participant's response (i.e., pressing "B" or "A"), however each excerpt continued to play through to completion after a response was provided. After each response, participants pressed the space bar to initiate the next trial.

In the test phase, 10 songs were presented. For each song there were seven probe-target line excerpts. Three of the target excerpts occurred before the probe line, and four target excerpts came after the probe excerpt. Thus there were seven trials per song, and hence 70 trials in the test phase. No accuracy feedback was provided in this phase.

The order of presentation of song type (popular vs. new) was counterbalanced across participants. Half of the participants listened to the popular song excerpts first, while the other half heard the new song excerpts first. Presentation order was further counterbalanced in the popular songs, with half of the participants hearing excerpts from the top of their chosen list of songs (i.e., the ones they were most familiar with) first, whereas the other half started with excerpts from songs on the bottom of their lists. The order of presentation of excerpts from the new songs was similarly counterbalanced across participants.

III. RESULTS

A. Reaction Times

For each participant, seven RTs were recorded for each song. RTs for each correct response to a target line were combined across songs to calculate a mean RT for each target type. The *Before* excerpts were Lines 2 and 3 (L2, L3) of Verse 1 (V1), while the *After* excerpts were Line 3 (L3) of Verse 2 (V2), and Lines 2, 3 and 4 (L2, L3, L4) of Verse 3 (V3). Mean RTs were calculated for the Popular songs, and for the New songs. RTs for L1 of V1 were not included in the calculation of mean RT. As indicated earlier, this line was included as a target line to test for a primacy effect in the examination of accuracy only, and also to balance the number of *Before* and *After* excerpts.

Two participants were identified as outliers, with their mean RTs being three or more standard deviations from the group mean RT. One participant was an outlier with respect to *After* RTs only, whereas the other participant was an outlier with respect to both *Before* and *After* RTs. Data from the remaining 48 participants were included in the RT analyses. Prior to the exclusion of these participants, the skewness values for the *Before* and *After* conditions were .725 and .813 respectively, and following exclusion the values were .074 and .451 respectively. Following exclusion of the outliers, the Shapiro-Wilks significance levels for the *Before* and *After* conditions were .488 and .105 respectively. The data therefore met the assumptions of normality for a repeated measures t test (N>30), and the assumptions of the analysis of variance (Tabachnick & Fidell, 2001, p.72).

A repeated measures t test was used to examine the effect of song type (i.e., popular vs. new songs) on overall mean RT. There was no significant effect, $t(47) = .409$, $p = .700$. As a result, all subsequent analyses ignored this variable. Mean RTs for each song type are presented in Table 2.

Table 2. Mean RT (*ms*) as a function of song type and target position. Values in parentheses are standard deviations.

	Song Type		
	Popular	New	Combined
Target Position			
Before			
V1L2	5003.21	4857.95	4961.15
	(1568.86)	(2203.24)	(1468.23)
V1L3	4954.92	5588.03	5110.69
	(1586.49)	(2196.63)	(1491.36)
After			
V2L3	4781.63	5663.77	4970.32
	(1618.13)	(2045.29)	(1542.84)
V3L2	4760.04	4377.15	4663.66
	(1696.86)	(2042.81)	(1591.76)
V3L3	4577.24	4277.58	4501.68
	(1620.28)	(1508.47)	(1480.65)
V3L4	4504.77	4433.31	4492.72
	(1537.11)	(1876.38)	(1456.72)

Note: Target position labels have the following convention – V1L2 = Verse 1, Line 2.

A repeated measures t test was used to compare mean RTs in the *Before* and *After* conditions. A significant difference between the two conditions was found, $t(47) = 2.63$, $p = .012$, with responses being faster in the *After* condition (M = 4657.10*ms*, SD = 1516.99*ms*) than in the *Before* condition (M = 5035.92*ms*, SD = 1479.80*ms*).

A series of analyses was conducted to test for the effect of song position while keeping line position constant. The first such analysis was a repeated measures t test that compared mean RT in V1L2 (a *Before* line) with mean RT in V3L2 (an *After* line). No significant difference was found, $t(47) = 1.77$, $p = .084$. The second analysis was a one-way repeated measures ANOVA comparing mean RTs for Line 3 of Verses 1, 2 and 3. A significant effect of Verse was found, $F(2,94) = 6.19$, $p<.05$. Post hoc tests were conducted using Tukey's HSD on the three possible pairwise comparisons. A significant difference was found between V1L3 and V3L3. As can be seen in Table 2, participants were faster at identifying Line 3 in the *After* condition than in the *Before* condition, but only in Verse 3 and not in Verse 2.

B. Accuracy

For each participant, the number of target lines responded to correctly were converted to Accuracy scores (%). For each participant Accuracy scores were calculated for the popular songs and the new songs. The minimum requirement for inclusion in the analysis was that a participant's Accuracy for each type of song was at least 50%. This was to ensure that the responses were better than chance. All participants scored equal to or better than 50%, so all responses were included in the analysis. However, data screening on the Accuracy scores indicated that there were two outliers with mean Accuracy scores more than three standard deviations below the mean. The data from these participants were eliminated from further analyses.

A repeated measures t test was performed to compare the overall Accuracy scores for the popular and new songs. No significant difference was found, $t(47) = .729$, $p = .493$, so subsequent analyses ignored the variable of song type. Average Accuracy scores for each target line for each song type are reported in Table 3.

Table 3. Mean Accuracy (%) as a function of song type and target position. Values in parentheses are standard deviations.

Target Position	Popular	Song Type New	Combined
Before			
V1L1	87.50 (15.25)	83.34 (25.96)	86.68 (13.42)
V1L2	75.00 (17.49)	80.21 (24.71)	76.04 (15.67)
V1L3	62.76 (20.22)	56.25 (42.05)	61.88 (19.75)
After			
V2L3	61.82 (21.65)	72.92 (34.14)	65.21 (16.50)
V3L2	70.31 (16.83)	72.92 (34.14)	71.46 (15.30)
V3L3	69.69 (19.42)	76.04 (32.60)	71.50 (16.37)
V3L4	83.59 (13.44)	81.25 (28.48)	83.13 (12.23)

Note: Target position labels have the following convention – V1L2 = Verse 1, Line 2.

A repeated measures t test was used to compare mean Accuracy scores in the *Before* and *After* conditions. No significant difference between the two conditions was found, $t(47) = 1.38$, $p = .174$ (*Before* $M = 74.87\%$, $SD = 16.28\%$; *After* $M = 72.82\%$, $SD = 15.10\%$).

A series of analyses was conducted to test for the effect of song position while keeping line position constant. The first such analysis was a repeated measures t test that compared mean Accuracy in V1L2 (a *Before* line) with mean Accuracy in V3L2 (an *After* line). No significant difference was found, $t(47) = 1.50$, $p = .140$. The second analysis was a one-way repeated measures ANOVA comparing mean Accuracy for Line 3 of Verses 1, 2 and 3. A significant effect of Verse was found, $F(2,94) = 3.30$, $p<.05$. Post hoc tests were conducted using Tukey's HSD on the three possible pairwise comparisons. A significant difference was found between V1L3 and V3L3. As can be seen in Table 2, participants were more accurate at identifying Line 3 in the *After* condition than in the *Before* condition, but only in Verse 3 and not in Verse 2.

IV. DISCUSSION

The current study tested the hypothesis that with respect to the direction of memory for music, 'forward is best' in both popular and new songs. Specifically it was hypothesised that RT would increase incrementally (i.e., line by line) the further forward participants searched through a song to identify the location of a target line in relation to a probe line.

The main findings of the study were: (1) there was no significant difference in the speed or accuracy of responding to popular and new songs; (2) RT for correctly identifying target lines that fell after the probe line were generally shorter than for those that occurred before the probe line, although this difference was not as predicted; (3) there was no overall difference in Accuracy between *Before* and *After* lines, however, participants were more accurate at identifying Line 3 in the later Verse 3 than in the earlier Verse 1.

The finding that there was no difference in mean RT between the popular and new songs supports the findings reported by Peretz, Gaudreau and Bonnell (1998) that memory for an unfamiliar song is reliable and accurate even after only a single playing of the song. The current finding also provides support for Schulkind's (1999) conclusion that those characteristics that support learning and memory for song, such as temporal information, are as effective in novel songs as they are in familiar songs.

The finding that mean RT in the *After* condition was faster than in the *Before* condition supported the general hypothesis of the current study that 'forward is best' in memory for music. This result is consistent with Wallace's (1994) assertion that the repetition of melody across verse facilitates recall of song lyrics to the extent that the lyrics towards the end of the song are recalled and recognised with more ease than the earlier lyrics of the song. However, when Line 3 RTs for each verse were compared, the only significant difference was between Verses 1 and 3. Mean RT in V2L3 was not significantly different to either of the other Line 3s, although it did fall in between the values for these two lines, as predicted by the forward bias hypothesis.

Overall, the slowest mean RT was for V1L3, as predicted, however the prediction that mean RT would increase incrementally with each line following the probe line was not supported by the results. In particular, V2L3 was predicted to have the fastest mean RT as it directly followed the probe line. However, performance on this line was the second slowest of all lines. A similar result was observed in each of the individual songs used in this study. This line elicited the slowest responses in the new condition, and the third slowest responses in the popular condition.

One possible explanation for this anomaly comes from the common observation that the more alike two stimuli are, the more difficult it is to distinguish one from the other. For instance, when people are asked to make mental comparisons of objects along some dimension, the closer the objects are on that dimension, the more time it takes to make the comparison, and the comparisons tend to be less accurate (Moyer, 1973; Potts, 1972). In this experiment, upon hearing the target line that comes immediately after the probe line, participants may recognise that the lines occur in close temporal proximity in the course of the song, but require extra time to decide which comes first. Conversely, if the two lines are further apart in the song the comparison is easier to make. This finding warrants further investigation. One possible method would be to record different pairs of consecutive lines within individual songs, and test whether the finding is consistent both within and across a variety of songs.

The fastest RT was recorded for V3L4, a finding that reflects a recency effect. Even with this line, and the line immediately following the target line (V2L3) discounted, the predicted order of RT was still not found. In Verse 3, RT for Line 3 was faster than for Line 2, the opposite to what was predicted. So, although the overall *Before*/*After* difference is consistent with participants searching for the target line in a forward direction from the probe line through their memories of each song, the fact that RT increased for the *After* lines as the distance from the probe line increased suggests that the search process is not as simple as originally hypothesised.

There was no difference in Accuracy scores for the popular and new songs. In contrast to the RT results, there was no difference in Accuracy between the *Before* and *After* conditions, which reflects the influence of a clear primacy for the first line of the songs. This finding echoes the strong primacy effect that Rubin (1977) found in his experiment on long-term memory for prose. There was also evidence of a recency effect in the current experiment, as the next most accurate score was for the last line of Verse 3. The pattern of Accuracy scores throughout the songs suggests that accuracy began at a high level (primacy effect), decreased until reaching the lowest level in the middle of the songs, then increased again by the end of Verse 2 until reaching the high score for the last line (recency effect).

A companion study was performed to try to replicate these findings and examine whether musical expertise has any influence on the results. Those with musical expertise have been found to possess superior recall for both text and melody presented separately and together when compared to those without this expertise (Ginsbourg & Sloboda, 2007; Kilgour, Jakobson & Cuddy 2000). The companion study broadly replicated the pattern of results in the current study, and indicated no effect of musical expertise.

V. CONCLUSIONS

The current study tested the hypothesis that in memory for music 'forward is best'. The results of the experiment provided partial support for this hypothesis in that a general bias towards searching through songs in a forward direction was found for both popular and new songs. Further, when responses on the equivalent line (L3) in different verses were compared, the pattern of results were as predicted by the hypothesis. This conclusion suggests that song memories do not have the random-access nature of digital memories (e.g., in computers), but retain the linear nature inherent in musical structures. This linear feature appears to remain even in highly familiar songs. However the specific prediction that RT would increase as the distance between the probe line and the target line increased was not supported. In particular, responses to the line that followed immediately after the probe line (V2L3) were generally the slowest and least accurate, suggesting that there are effects related to the structural differences between lines in a song that may overlay the forward bias of memory for songs.

REFERENCES

Bartlett, J.C. & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology*, *95*, 551-560.

Dowling, W. (1973). Rhythmic groups and subjective chunks in memories for melodies. *Perception and Psychophysics*, *14*, 37-40.

Frances, R. (1988). *The perception of music.* (W.J.Dowling Trans.). Hillsdale, NJ: Erlbaum. (Original work published in 1958.)

Gérard, C. & Auxiette, C. (1992). The processing of musical prosody by musical and non-musical children. *Music Perception*, *10*, 93-126.

Ginsborg, J. & Sloboda, J. (2007). Singers' recall for the words and melody of a new, unaccompanied song. *Psychology of Music*, *35*, 421-440.

Kilgour, A., Jakobson, L. & Cuddy, L. (2000). Music training and rate of presentation as mediators of text and song recall. *Memory and Cognition*, *28*, 700-710.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Moyer, R.S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception and Psychophysics*, *13*, 180-184.

Peretz, I., Gaudreau, D. & Bonnell, A. (1998). Exposure effects on music preference and recognition. *Memory and Cognition*, *26*, 884-902.

Potts, G.R. (1972). Information processing strategies used in the encoding of linear orderings. *Journal of Verbal Learning and Verbal Behaviour*, *11*, 727-740.

Rubin, D.C. (1977). Very long-term memory for prose and verse. *Journal of Verbal Learning and Verbal Behaviour*, *16*, 611-626.

Schulkind, M.D. (1999). Long-term memory for temporal structure: Evidence from the identification of well-known and novel songs. *Memory and Cognition*, *27*, 896-906.

Serafine, M.L., Crowder, R.G. & Repp, B.H. (1984). Integration of melody and text in memory for songs. *Cognition*, *16*, 285-303.

Serafine, M.L., Davidson, J., Crowder, R.G. & Repp, B.H. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language*, *25*, 125-135.

Tabachnick, B.G. & Fidell, L.S. (2001). *Using multivariate statistics* (4^{th} ed.). New York: Harper Collins.

Tan, S. & Speckman, M. (2005). Listeners' judgements of the musical unity of structurally altered and intact musical compositions. *Psychology of Music*, *33*, 133-153.

Wallace, W. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*, 1471-1485.

Wallace, W. & Rubin, D. (1991). Characteristics and constraints in ballads and their effect on memory. *Discourse Processes*, *14*, 181-202.