Musical Accents and Memory for Words
Thomas Ting¹ and William Forde Thompson²
Department of Psychology, Macquarie University, Australia

¹Thomas.Ting@students.mq.edu.au, ²Bill.Thompson@mq.edu.au

ABSTRACT
In this study, we examined the effect of background music on reading, focusing on memory for words that are read concurrently with musical accents. Can musical accents enhance memory for words in the same way that visual accents (underlining, highlighting) draw attention to words and hence increase memory for them? Forty undergraduate psychology students were presented sentences one word at a time on a computer screen. Each word was accompanied by a piano tone such that the sequence of tones outlined a brief melody with one note being musically accentuated. Melodic accents had increased intensity, duration and pitch height. There were three music conditions. In the first two, musical accents were either congruent (aligned) or incongruent (not aligned) with a target word. In the third condition there was no accompanying music. The target words were either visually emphasized (bolded) during the exposure phase or not. Surprisingly, recall was better when a musical accent was incongruent with a target word compared to when the accent was congruent or when there was no music at all. There was no significant effect of bolding target words in the exposure phase. The results suggest that background music enhances coding of words during reading, but only for words that do not coincide with strong musical accents. A cost-benefit trade-off model is suggested, where prominent musical accents may compete for attention, eliminating the potential benefits of accompanying music for verbal memory.

I. INTRODUCTION
Music is an important part of everyday life, often accompanying routine activities such as shopping, driving and studying. The widespread use of music across a wide range of daily contexts suggests the need for research on the cognitive consequences of background music. Empirical research has confirmed that background music can influence many human behaviours and cognitive abilities. For example, people take longer to shop while listening to unfamiliar music compared to when listening to familiar music (Yalch & Spangenberg, 2000). Radio advertisements are recalled more easily when accompanied by slow rather than fast music (Oakes & North, 2006). As background music increases in speed, so does simulated driving speed and perceived speed estimates (Brodsky, 2001). In films, background music enhances memory for scenes where auditory and emotional cues are related in mood, compared to when they are of differing moods (Boltz, 2004).

A. Background Music and Reading
Studying with background music is common practice for many students. Fendrick (1937) reported one of the earliest studies of the effects of background music on reading. Reading efficiency was examined among individuals who read in the presence of either lively “semi-classical” music or without any sound. Results showed that the background music was distracting: reading ability was better with no sound than with music. In a subsequent study, Henderson, Crews and Barlow (1945) used the Nelson-Denny Reading Test to evaluate reading efficiency in the presence of either popular or classical background music. Results illustrated that the classical music was not distracting whereas the popular music resulted in significant distraction effects.

Freeburne and Fleischer (1952) further explored the effect of musical genre on reading speed and comprehension. Subjects were required to complete a reading comprehension task while listening to classical, semi-classical, popular, or jazz music in the background. They observed no significant differences between musical genres on comprehension and number of lines read.

Most recently, Thompson, Schellenberg and Letnic (in press) assigned a challenging reading comprehension task to students while they listened to classical background music. The same musical piece was played in all conditions, but the overall tempo and intensity of the track was manipulated to create four background music conditions: slow-soft, slow-loud, fast-soft, fast-loud. Compared to baseline (reading in silence), reading comprehension was impaired only when the background music was fast and loud. For all other conditions, there was no significant effect of background music on reading comprehension.

B. Syntax in Language and Music
There are important similarities between the processing of music and language as “both domains create hierarchically organised sequences based on abstract structural categories” (Patel, 2010, p. 4). Both music and language involve integrating new events (words or tones) with past events, segmentation, phrase boundaries, and meaningful points of stress or accent. Because of these similarities, there has been speculation that the processing of music and language may share cognitive processes.

One plausible process that may be shared between these two domains is syntactic processing. Syntax involves structural principles that govern whether a sequence of linguistic material makes sense. In language, syntax can constitute combinations of words, phrases and sentences (Patel, 2003). While syntax is a large component of linguistics, it is not confined to language. In music, syntax is evident through the use of chord progressions, time signatures, keys and phrases (Patel, 2003). Thus, music and language are likely to share syntactic processes because both consist of temporally structured acoustic signals.

C. Overlap in Syntax Processing
Patel’s (2003) “shared syntactic integration resource hypothesis” (SSIRH) argues that there is overlap in syntactic processing between language and music. More specifically, the SSIRH proposes the existence of a region in the brain that is responsible for processing of syntax in
language and music (see also Patel, 2009). Because the SSIRH suggests that both domains recruit the same brain processes, it also predicts that interference effects should arise whenever a task requires both musical and linguistic syntactic processing.

Slevc et al. (2009) tested this very hypothesis. Participants read complex sentences (garden path sentences) in a self-paced manner by clicking on a mouse to progress to subsequent phrases of the sentence. As each segment of the sentence was presented, a musical chord was also presented. The musical syntax manipulation involved the occurrence of an unexpected chord. When unexpected accents occurred, self-paced reading time slowed down. The authors argued that the processing of unexpected musical chords taxed the same resources required to process the garden path sentences, thereby interfering with the rate of processing (see also, Fedorenko, Patel, Casasanto, Winawer and Gibson, 2009).

Koelsch (2005) reported neuroscientific evidence that music and language are processed by similar neural resources. Utilising magnetoencephalography (MEG), he observed that the processing of musically irregular chords activates the inferior frontal cortex. This area in the left hemisphere is called Broca’s area and is associated with the processing of linguistic syntax. Similarly, Sammler, Koelsch, Ball et al. (2009) examined EEG responses to syntactic violations in chord progressions and sentences. Their results again support Patel’s (2003) model and suggest that the bilateral superior temporal and left inferior frontal lobes may be sites that are recruited by both music and language. Other studies using fMRI have shown that the processing of unexpected chords activates both Broca’s area and the homotopic area of the right hemisphere (Koelsch, Fritz, Schulze, Alsop & Schlaug, 2005; Koelsch, Gunter, Cramon, Zysset, Lohmann & Friederici 2002; Tillman, Janata & Bharucha, 2003).

D. Accents and Memory

Patel (1998) emphasised that both music and language are dependent on memory for structural elements. The elements of language include words, whereas the elements of music include chords and tones (Patel, 2010). In music, some events are often made to stand out amongst others through changes in duration, intensity, pitch height, and timbre (Snyder, 2001). Musical moments that are emphasised in this manner are labelled as accent. Language has its own representations of points of stress. Auditory linguistic accents are similar in that they also utilise changes in duration and loudness. Visual linguistic accents such as the use of bold, capitalisation, underscoring or italics indicate points of stress as well. In all cases, accents facilitate syntactic processing, highlighting the importance of the accented information. For example, the sentence “John gave MARY the purse” implies that Mary received the purse and not someone else. When the accent is applied differently as in, “John gave Mary the PURSE”, it suggests that the object of a purse was given, rather than another object.

Changes in pitch can also lead to changes in meaning (Koelsch, 2005). Pitch changes can help to differentiate questions from statements, and can assist in the perception of phrase boundaries. As such, accents and pitch changes play a critical role in the accurate perception of syntactic structure in both music and language.

E. The Present Study

In the current study, we manipulated musical accents in order to explore the syntactical relationship between music and language. Our goal was to determine whether musical accents can enhance memory for words in the same way that musical underscoring can draw attention to events in a film. For example, in the film “Jaws” the viewers’ attention to an approaching shark is enhanced by the introduction of the menacing theme music, making the visual image more compelling and memorable. Similarly, in the stabbing shower scene from the movie “Psycho”, the impact of the murder is increased dramatically by the introduction of short high-pitched musical “stings” in the underscoring. These examples illustrate that musical accompaniment can draw attention towards accompanying media, implying that musical underscoring is not “informationally encapsulated” (Fodor, 1983) but can interact with other media being processed within the same time window.

When music and other media are processed simultaneously, accents in one channel (e.g., music) are treated equivalently to accents in the accompanying channel (e.g., reading material). According to this “accent-sharing hypothesis” the effect should also occur when music and reading materials are combined. An accent in the music should draw attention towards all events in that temporal window, thereby enhancing the encoding of those events. It is therefore hypothesised that when melodic accents coincide with words, they may be encoded as verbally accented. This “spill over” accenting effect should be reflected in enhanced memory for those words. When musical accents fall on target words, recall of target words and recall reaction time should be better than when the accents fall on non-targets and when there is no sound.

II. METHOD

Forty (9 males, 31 females) undergraduate psychology students from Macquarie university took part in the study, with age ranging from 18 to 38 years (M = 20.68). They were given partial course credit in exchange for their participation. All participants reported having normal hearing ability.

Seventy-five sentences were used, and were either original or adapted from Bloom and Fischler’s (1980) list of sentence completion norms. The length of the sentences ranged from five to eleven words. Within each sentence was one target word. The sentences were presented in two conditions. In one condition the target words were capitalized; in the other condition the target words were not capitalised. In addition to the experimental sentences, 25 grammatically incorrect sentences (grammatical foils) were also used. Melodies were created for 2/3 of the experimental sentences. Melodies were composed in a major key with starting notes varying between F4, F#4, G4 and G#4. All melodies followed an ascending then descending contour. The start and end notes were always the tonic (the first scale degree). The melodies were presented in a piano timbre. The number of notes in the melody was identical to the number of words in the corresponding sentence it accompanied. One
note within each melody was musically accented. A combination of four musical aspects was used to create the melodic accents. The accented notes were:

- The highest pitch in the melody
- A point of contour change
- Louder than other note (20dB higher in amplitude)
- Longer than other notes (300ms against 150ms)

Figure 1 provides an example of the verbal and melodic materials. Experimental sentences were presented in three conditions with twenty-five sentences per condition:

- **Target accented**: Sentences were accompanied by a melody with an accent occurring on a target word.
- **Non-target accented**: Sentences were accompanied by a melody with an accent occurring on a non-target word. Accents appeared one, two or three words before or after the target word.
- **No sound**: Sentences were presented without melodies.

Five examples for each of the above conditions are shown in the Appendix. Out of the 25 grammatical foils, 16 were presented along with melodies while 9 were not given sound. In total, 100 sentences were presented, and are available by request from the first author. The experiment was split into two sections. In the first, participants were instructed to judge the grammatical correctness of each sentence. The sentences were presented one word at a time in the centre of the screen. Each word was presented for a length of 300ms using the font Arial at size 48. Presentation of each word coincided with the onset of the musical note in all conditions except the no sound condition. At the end of every sentence the following question appeared on the screen: “Was this sentence grammatical?” Subjects were told to respond by pressing one of two keys to indicate “yes” or “no”. If a response was correct, the word “CORRECT” appeared for a short duration before the next trial. Similarly, if a response was incorrect, the word “INCORRECT” appeared briefly. The sentences were presented in a random order and began with a fixation cross. Participants were not instructed to respond quickly or accurately in this section.

In the second stage of the experiment the subjects were given a surprise memory test immediately after the first stage concluded. The target words from the grammatically correct sentences presented in the exposure phase were presented along with an equal number of foils. Foils were similar to target words and would have fit the sentence if they replaced the real target. The words appeared one at a time in a random order in the centre of the screen. A total of 150 words were presented (75 targets and 75 foils). The participants were told to indicate whether they had seen the word in the first part of the experiment by pressing the appropriate button. They were instructed to respond as quickly and as accurately as possible. No feedback was given in the recall task.

For half of the participants, the target word in the grammatical judgement task and all the words in the recall task (target and foils) were presented in upper case. All the remaining words within the sentences were in lower case. For the other half of the participants, there were no words that were completely capitalised. The capitalisation of words was not mentioned to participants. Once the recall task finished, participants were asked to complete a short questionnaire regarding basic demographic information.

Data from participants who correctly identified less than 66% of the grammatical foils in the grammatical judgment task were rejected and were not further analysed. For all the accepted participants, two types of responses were recorded for both tasks: reaction time (RT) and accuracy. The reaction time data from both the grammatical judgement task and recall task were analysed using a mixed analysis of variance (ANOVA) with the between subject factor as FONT (upper case vs. lower case) and within subject factor as CONDITION (target accented vs. non-target accented vs. no sound). The recall accuracy data was analysed using a mixed ANOVA with between subject factor as FONT (upper case vs. lower case) and within subject factor as CONDITION (target accented vs. non-target accented vs. no sound) on hit minus false alarm rates. An alpha of .05 was set as the criterion for statistical significance.

### III. RESULTS

Of the 40 participants, 30 valid cases were obtained. 17 of the valid cases were from the upper case group, while 13 were from the lower case group. It was predicted that when musical accents coincide with target words, memory for those words would be improved. Additionally, when musical accents are mismatched with target words, memory would be reduced for those words.

#### A. Effect of Music Condition on Recall Accuracy

There was a significant main effect of music condition \(F(2,56) = 3.309, p = 0.044\), although not in the direction hypothesised. As shown in Table 1, recall was higher when a non-target was accented compared to when there was no sound \(F(1,28) = 5.324, p = 0.029\) and when compared to the target being accented \(F(1,28) = 4.396, p = 0.045\). Although recall rates for targets may imply guessing, it should be noted that these rates were all significantly higher than the false alarm rate for “foils” that were not in the exposure phase (M = .23, SD = .15). When presented with unfamiliar words, participants usually indicated they had not seen them.

#### Table 1. Mean and standard deviations of target recall scores.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Accented</td>
<td>0.48</td>
<td>0.17</td>
</tr>
<tr>
<td>Non Target</td>
<td>0.53</td>
<td>0.18</td>
</tr>
<tr>
<td>Accented</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Sound</td>
<td>0.48</td>
<td>0.18</td>
</tr>
</tbody>
</table>

#### B. Effect of Music Condition on Reaction Time

Between the three critical conditions there was no significant difference in reaction time on the grammatical task \(F(2, 56) = 1.721, p = 0.188\). There was also no significant effect of music condition on recall reaction time \(F(1,485,
41.590) = 1.855, \ p = 0.177]. Overall, it appears that music condition has no effect on reaction time in both recall and grammatical tasks. Interestingly though, there appears to be a trend in the grammatical judgement task where sentences with musically accented targets were responded to faster than when there was no sound, although analysis yielded results just outside of statistical significance [\ F(1,28) = 3.737, \ p = 0.063].

C. Effect of Font

There was no significant effect of text capitalisation (upper case vs. lower case) on recall accuracy [\ F(1,28) = 0.215, \ p = 0.646]. There was also no significant effect of capitalisation on recall reaction time [\ F(1,28) = 0.945, \ p = 0.339] or grammatical reaction time [\ F(1,28) = 2.043, \ p = 0.164].

IV. DISCUSSION

The results of this study indicate that memory of words may be significantly influenced by the presence of background music, but only if musical accents do not coincide with the words that are to be remembered. We expected that recall of target words would be improved with a coinciding melodic accent. Surprisingly, however, recall was only better when musical accents did not correspond with target words; when they coincided with the target words, the beneficial effects of music disappeared. One interpretation of this result is that melodic accents were a source of distraction, such that words coinciding with an accent are more poorly remembered than words not coinciding with a melodic accent. Font had no effect on recall accuracy, recall reaction time, or grammatical reaction time. We observed no reliable effects of the conditions on reaction time.

Two explanations are considered for the improvement in recall when a melodic accent fell on a non-target word. First, background music may have generally facilitated target recall, increasing memory for all words in the sentence except for words that accompanied a melodic accent. Second, music may have resulted in forward and backward priming effects, thereby improving recall for words that occurred immediately before or after the melodic accent. In either case, melodic accents may have been a source of distraction such that words coinciding with that accent were more poorly remembered than other words in the stimuli. These interpretations will be discussed in turn.

A. Are Musical Accents a Source of Distraction?

Memory for target words was worse when they coincided with melodic accents, suggesting that those accents were a source of distraction. Whitely (1934) also examined the effect of background music on memory. On the whole, performance was slightly lower when background music was present compared to when there was no sound. Unlike Whitely’s (1934) study, the present study used simple, rather than complex, auditory material. Our study suggests that individual musical accents alone can be a source of distraction.

B. Effects of Mood and Arousal

Background music enhanced memory for target words, as long as those words were not accompanied by an accented musical event. This enhancement effect may arise from increases in arousal and mood, which are associated with music listening and are known to enhance cognitive function (Thompson, Schellenberg, & Husain, 2001). Thompson et al. (2001) observed that increased arousal induced by energetic music facilitates subsequent performance on nonmusical tasks. On the other hand, Whitely (1934) reported that neither positive nor negative music had any effect on memory for accompanying verbal material. Taken together, these findings suggest that arousal and mood may not influence all cognitive tasks, and may not influence memory. Indeed, individuals exposed to Mozart’s music performed no differently to those exposed to a recording of rain or silence before a backwards digit span task, a measure of memory (Steele, Ball, & Runk, 1997).

In the present experiment, we observed a significant effect of music on memory, but only for words that are not accompanied by a melodic accent. There are important differences between our study and those of Steele et al.’s (1997) and Whitely’s (1934). First, we did not ask participants to engage in explicit memorisation, which is a conscious and active process. That is, they were not forewarned of the upcoming recall task. Second, in our study music was presented during the memorisation process instead of beforehand.

C. Forward and Backward Priming

Background music led to better memory for target words as long as the melodic accent did not coincide with those target words. The effect may be explained as an outcome of arousal, as described above. However, because accents were present in both music conditions, the effect may also be explained as an indirect outcome of melodic accent that coincided with a surrounding (non-target) word. That is, the musical accent may have primed surrounding words, but only those that did not coincide with the accent (and were therefore not subject to distraction) benefited from this priming effect. A typical priming study involves the presentation of a priming stimulus followed by a target word (Hermans, Houwer, & Eelen, 2001). For example, the time needed to evaluate a target word as positive or negative is faster when both the prime and target are congruent in valence. Fockenberg, Koole and Semin (2006) have also demonstrated a backward priming effect that occurs when a prime appears after a target word has appeared. Therefore, both forward and backward priming effects are evident in language (see also Meyer and Schwanefeldt, 1971).

With evidence supporting an overlap in the processing of meaning in music and language (Koelsch, 2005; Steinbeis & Koelsch, 2005), it is possible that the musical accents highlighted the semantic importance of the accompanying linguistic material. In combination with forward and backward priming effects, words surrounding a musically accented word may have been encoded more deeply in memory. When the accent coincided with a target word, a distraction effect by the accent itself may have countered any positive effect based on priming, yielding no effect. Thus, increased performance in the ‘non-target accented’ condition may have resulted from a combination of forward and backward priming effects. Controlled studies are needed to investigate this possibility.

D. Cost-Benefit Trade-Off Model

It has been proposed that there are both costs and benefits of background music for concurrent cognitive activity
(Thompson, Schellenberg & Letnic, in press). Benefits include heightened arousal, positive mood, and an "underscoring effect", such as that capitalised in film. The underscoring effect refers to the way in which music is able to emphasise accompanying media in the same way that a coloured highlighter does. Costs include distraction effects and capacity limit effects: when music draws attention to itself, it can distract a person from a primary task, such as navigation (driving while listening to music), reading comprehension, or memory for read materials. When there is a musical accent placed directly on a target word, this draws attention away from the target word, thereby undermining the potential positive effects of the music. Put differently, an "accent-distraction effect" may have cancelled out any potential benefit from underscoring. It was observed that there were benefits of music for concurrent cognitive activity when a non-target word was accented. When there was an accent on the target word the benefits and costs cancelled each other out, resulting in performance that was no different from the "no music" condition.

E. Implications for Future Research

The current study highlights a number of issues relevant to the effect of musical accents on memory. Accenting had a significant effect on memory for accompanying verbal material. However, the musical accent was composed in such a way that it may have drawn too much attention to the melodic event, creating a distraction effect. It is possible that the introduction of more subtle accents would lead to enhanced memory for accompanying words. It should also be noted that background music is often quite complex, whereas the accompanying melodies used in this study were simple. The simple melodies in the present experiment may have facilitated memory. On the other hand, studies involving complex auditory material (Freeburne & Fleischer, 1952; Henderson et al., 1945; Fendrick, 1937) usually do not control for a number of variables such as timbre, tempo and dynamic contrasts, and results across investigations have been inconclusive.

We did not ask participants about their study habits. Study habits should be investigated in order to determine if there are differences in memory performance for individuals who regularly study with background music and those who do not. Kotsopoulou and Hallam (2010) surveyed school and university students about their study behaviours in relation to background music. The questionnaire study explored the kinds of tasks performed, perceived effects of the music, types of music used, and factors that influenced the decision to start and stop music. The results indicate that students did not play music during intense study, rarely played music when memorising material, and most often played music while thinking and writing, suggesting that students were aware of the impact that music may have on learning. It was also reported that older students were more likely to turn off music when they felt it was distracting, implying that they are more conscious of the distracting effects of music. Alternatively, they may require greater focus because of the increased difficulty of tasks associated with higher levels in education (Kotsapoulou & Hallam, 2010).

The current study involved simple melodies when the effects of music on memory, and our stimuli were constructed to maximise experimental control. It may be argued that most background music used while studying is contemporary popular music. Indeed, popular music is the most common genre reported for use as an accompaniment to other activities. Whether musical accents often interact with verbal processing under normal circumstances is unclear. However, our goal was not to survey effects that occur in naturalistic settings, but to elucidate a process that may be shared by music and language.

One of the most common features of popular music is the presence of lyrics – a factor not considered here. Söqvist, Halin and Hygge (2010) examined the effect of task-irrelevant speech on reading comprehension, finding that irrelevant speech impeded comprehension. Interestingly, individuals who had higher working memory capacity were less distracted by irrelevant speech. Their results suggest that people who can immediately suppress unrelated material perform better at reading comprehension (Söqvist et al., 2010).

Although there was no significant difference in reaction time between conditions as predicted from Slevc et al.’s (2009) study, it is possible we had insufficient statistical power to elicit such an effect. We did observe a non-significant trend in which participants required less time to determine the grammatical correctness of a sentence when a musical accent fell on a target word than for other conditions. Replicating the current investigation with a larger sample size may clarify whether this trend is reliable.

V. CONCLUSION

The current investigation manipulated musical accents to explore the connection between music and language. Contrary to predictions, recall for target words was better when non-target words were musically accented compared to when target words were accompanied by a musical accent or when there was no accompaniment. The pattern of results may be explained by a cost-benefit trade-off model of the effects of music listening on concurrent cognitive activity. The presence of background music may confer a general benefit to memory for accompanying verbal material by underscoring that material. Alternatively, the musical accents may have highlighted surrounding words as important. However, when music captures too much attention, as would occur during melodic accents, there is a cost to accompanying cognitive activity in the form of distraction. Such a cost-benefit trade-off model explains why memory for verbal material was enhanced for the target-unaccented music condition, but not for the target-accented music condition.

ACKNOWLEDGMENT

We thank Adam Bentvelzen, Alex Chilvers, Catherine Greentree, Lincoln Colling, Vijay Marimuthu, Varghese Peter, and Lena Quinto for assistance and critical feedback.

REFERENCES


**APPENDIX**

Five examples of stimuli used in each of the three conditions. For music conditions, the target word (bolded) either coincided or did not coincide with a musically accented note (bolded).

**Condition 1: Target Accented**

1. Jack waited until the last **minute** to place all his bets.

2. The leg was **infected** quite severely.

3. The sudden **scream** surprised them.

4. The entrance to the **apartment** was blocked by a branch.

5. The cast thought the **script** was not great.

**Condition 2: Non-target Accented**

1. Not joining the **army** was Larry’s choice.

2. The person deserves our **gratitude** for identifying the thief.

3. The dirt from her **skirt** was removed.

4. Rick cleans his **wardrobe** every fortnight.

5. The quiet **hill** is where Ken built his shelter.

**Condition 3: No Sound**

1. Wally was too **poor** to buy another whiskey.

2. A slice of **tomato** made the sandwich memorable.

3. Paul was **upset** at being embarrassed.

4. They tightened the **cords** because the sail became loose.

5. The intense pain in his **back** stunned the senator.