Edinburgh Napier UNIVERSITY

# Is speech special? The impact of familiarity with irrelevant background sound on working memory

Nicola M$^{c}$Corkell

Supervised by: Alex Willis

April 2012

# Is speech special? The impact of familiarity with irrelevant background sound on working memory


## ABSTRACT

Working memory is a limited capacity storage system where information is retained and manipulated for a brief period of time. It is susceptible, however, to interference from corresponding sensory stimuli. Irrelevant background sound is consistently demonstrated to detrimentally interfere with the serial recall of visually presented items, however there remains debate for which properties of audio elicit the effect. The current study assessed the impact of familiarity with background music, along with speech, using variations of a familiar and unfamiliar song. A sample of 24 undergraduate students (11 males, aged *M* = 23, *SD* = 3.61) were opportunistically recruited. Six conditions were employed; silence, white noise, original track, otherwise equivalent instrumental music, recorded spoken lyrics and spoken lyrics over instrumental music. All participants were exposed to each condition in a pseudo-random order, while completing serial recall tasks. A significant main effect was found for sound (*p* <.01), and familiarity (*p* <.01), while speech over the instrumental music elicited the worst recall performance. The results suggest that fluctuation in acoustic properties is not the sole determinant of the irrelevant sound effect, but familiarity and speech may play a critical role in the degree of interference that unattended auditory stimulus provokes. Extensive analysis of which features in speech, and differing levels of familiarity, is required in order to reach definitive conclusions.



| KEY WORDS: | WORKING MEMORY | IRRELEVANT SOUND | INTERFERENCE | SPEECH | FAMILIARITY |
|---|---|---|---|---|---|

## 1.0 Introduction

Memory is essential for human functioning; we use it every day, even without conscious awareness. Without memory you would not be able to read nor comprehend this text, converse with peers, or even recognise your loved ones. But why is it that we sometimes find ourselves forgetting information? What is it that causes our memories to fail for things such as remembering someone's name we have just met, or a phone number which we read a few moments earlier? Questions such as these have been asked by many researchers, resulting in the production of various conflicting theories, each attempting to provide knowledge and understanding of how human memory works and what processes are involved.

One memory form which has been proposed is working memory, which refers to "a limited capacity temporary storage system that underpins complex human thought" (Baddeley, 2007, pp.6-7). This differs from short-term memory, as it is not simply the temporary storage of just encountered material, but also involves the manipulation of such material in mental space (Miyake, 2001). This higher cognitive functioning exemplifies our competence for remembering what we have just perceived, while thinking about it in relation to what we already know (Baddeley, 1986; Logie, 1996). However, this process is susceptible to interference from particular sensory stimuli or concurrent activities, although which inputs or activities interfere, and why they do so, continues to be a topic with substantial research and debate.

When considering the visual world, it is possible to omit or avoid unwanted inputs through the closure of ones' eyes or turning of ones' head, however, the auditory organ does not come equipped with such an equivalent. There is no natural mechanism available to exclude unwanted auditory stimulus, and so the material is inescapable. A phenomenon which has received an extensive degree of research, concerning the interference audio may cause, is the Irrelevant Sound Effect (ISE). It is believed that short-term serial recall of visually presented items is significantly influenced by interfering stimuli such as background noise (Baddeley, 1986; Colle & Welsh, 1976; Ellermeier & Zimmer, 1997; Jones & Macken, 1993, 1995; Salamé & Baddeley, 1982, 1989). It has been proposed that although the irrelevant material is arguably unattended, the level of interference on memory may indicate the nature and extent of pre-attentive processing of sound (Jones, Alford, Bridges, Tremblay & Macken, 1999).

How and why auditory stimulus affects serial recall has widely been debated throughout the literature, with little consensus on the mechanisms, or the particular characteristics of sound, which are entailed in causing the effect. Currently, there is debate concerning the effects of familiarity with irrelevant background noise, and whether vocal, or non-vocal music produces a greater effect. However, there appears to be no published literature at present which compares such conditions with that of familiar and unfamiliar speech. Thus, investigating this area will provide a better understanding of which precise properties of irrelevant background noise elicit the ISE. Therefore when applying this laboratory phenomenon to everyday life, the detrimental effects that background noise has on situations which rely on serial recall; such as remembering a phone number, or learning the sequence to a mathematical equation, can be better understood. Thus prevention measures can be implemented to limit the negative impact on performance.

## 2.0 Literature review

This section will initially define the key theories and ideas significant to the project, and will continue to review relevant contemporary literature in the discipline.

## 2.1 Working Memory Model

The working memory model, proposed by Baddeley and Hitch (1974), initially comprised three-components; an attentional control system known as the central executive, and two subsidiary storage systems; the phonological loop and the visuospatial sketchpad. The episodic buffer was later introduced as an interface between the three subsystems and long-term memory (Baddeley, 2000). The visuospatial sketchpad is assumed to retain and manipulate visual and spatial images, which can be useful for spatial problem solving, while the phonological loop is believed to hold and manipulate verbal information.

Within the phonological loop, spoken stimulus is assumed to have automatic access, however visually presented material requires phonological recoding (conversion of grapheme-to-phoneme), through subvocal articulation, to enable access to the phonological store. The storage of such material is believed to be limited, occur for a short period of time and be dependent on rehearsal (Baddeley, 2003), where rehearsal relies on the overt or covert vocalisation of the stimuli, one after another. Without this, it is believed that the memory trace would fade (Baddeley, 2007), however, when there are other processes occurring parallel to rehearsal, the recollection of the stimuli may be hindered.

## 2.2 Irrelevant Sound Effect

Colle and Welsh (1976) were the first to demonstrate the disruption interfering auditory stimuli can cause, through the use of spoken text in an unfamiliar language. Participants took part in a serial order memory task of visually presented letters, and although they were instructed to ignore the irrelevant speech, a considerable hindrance to their recall was evident due to the irrelevant background noise. Following this finding, many researchers have argued why this interference may arise, with several disputing theories proposing cognitive processes which may provide an explanation.

### 2.2.1 Interference-by-content

It was originally believed that only verbal auditory stimuli caused interference with serial recall, through an interference-by-content approach (Baddeley, 1986; Salamé & Baddeley, 1982). An auditory filter was thought to protect the phonological store, by preventing sounds which were non-speech-like from being transmitted (Baddeley, 1986). It was argued that disruption occurred due to the correspondence of characteristics in the to-be-remembered and to-be-ignored items, such as phonological or semantic similarity. Baddeley (1986) claimed the irrelevant speech effect occurred as a passive side-effect of the to-be-ignored verbal auditory information gaining automatic access to the limited capacity phonological loop, resulting in the same representational space as the to-be-remembered items being utilized. It has been found that the effect occurs not only when irrelevant noise is in an unfamiliar language (Baddeley & Salamé 1986; Colle & Welsh 1976; Salamé & Baddeley 1986, 1989), but also when speech is nonsense syllables (Salamé &

Baddeley, 1982). Salamé and Baddeley (1982) found the effect occurred no more when background noise was a sequence of irregular digits than for items containing the same phonemes of the digits but in an altered order (e.g. tun-woo rather than one-two). Furthermore, there appears to be many studies which demonstrate that phonological and/or semantic resemblance between the two materials show no additional interference to recall (Bridges & Jones, 1996; Jones & Macken, 1995a; LeCompte & Shaibe, 1997), questioning the strength of the initial proposal of the working memory model.

There have been some surprising features discovered in relation to the extent to which background noise disrupts performance. It was observed that the effect does not depend on the volume of sound (Colle & Welsh, 1976), nor does it lessen over time, either across many days or within the experimental period (Hellbrück, Namba, & Kuwano, 1996; Jones, Macken, & Mosdell, 1997). Additional doubt was cast over the validity of the initial working memory model by the discovery that short-term serial recall was also affected by music, thus was not exclusive to verbal stimuli (Salamé & Baddeley, 1989). Furthermore, noise stimulus, which fluctuated in intensity or pitch, caused the effect to arise (Salamé, 1990). The auditory filter hypothesis could, however, continue to support these results, given an additional assumption; it would be unlikely that a filter which permitted speech would be a perfect filter, consequently, other sounds which were speech-like in nature, would also be admitted (Larsen, Baddeley & Andrade, 2000).

### 2.2.2 Interference-by-process

As a result of such findings, Jones and his colleagues have extensively explored which kinds of stimuli, other than speech, may produce the irrelevant sound effect. It has been shown that the repetition of a single speech sound does not generate the effect, whereas a list of diverse speech sounds does (Jones & Macken, 1993; Jones, Madden & Miles, 1992; Jones, Macken & Murray, 1993). It was also demonstrated that pure tones which change in frequency cause interference (Jones and Macken, 1995b; Jones *et al*., 1993), thus indicating speech is neither essential, nor adequate, to generate the irrelevant sound effect, consequently disputing Baddeley's (1986) initial interference-by-content theory. Instead, Jones *et al*. (1992) suggest an interference-by-process approach, proposing the changing state hypothesis. It was identified that serial recall is solely disrupted through the fluctuation of acoustic properties of any irrelevant auditory stimulus, whether it is speech or non-speech sound. The lack of semantic and phonological similarity effect between to-be-remembered and to-be-ignored stimuli, which is one of the core findings the phonological loop hypothesis cannot account for, can however, be explained by this model. The crucial feature identified here is the relationship *between* the items within a sequence, not the items themselves.

## 2.3 Seriation

It has been found that tasks which do not rely on seriation, such as category recall and the missing item task, experience minor, if any, hindrance due to irrelevant background noise (Beaman & Jones, 1997; Perham, Banbury & Jones, 2007). Hence it is this seriation feature, in the serial recall task, which is crucial for the effect and shows most importance. This is because it is claimed to underlie various everyday activities, such as mental arithmetic (Banbury & Berry, 1998) and free

recall (Beaman & Jones, 1998), together with being inextricably related to language learning and production (e.g. Gupta & MacWhinney, 1997; Jones, Hughes & Macken, 2006; Perham, Marsh & Jones, 2009; Schweppe & Rummer, 2007). Damage to recall then develops from the conflict of processing two parallel sources of order stimulus; one from the intended use of the seriation, to retain the sequence of the to-be-remembered material, and the second from the preattentive processing of items in the acoustically variable irrelevant sound.

### 2.4 Alternative Theories

Jones and Macken (1995a) proposed the Object-Orientated Episodic Record (O-OER) hypothesis, where a specialised mechanism is suggested to justify why seriation tasks are required to exhibit the ISE, unlike Baddeley's (1986) phonological loop hypothesis. It is suggested that both auditory and visual stimuli are denoted on a metaphorical blackboard within short-term memory, where they are seen as 'objects'. Similar to Baddeley's phonological loop, auditory stimulus (e.g. the irrelevant sound) has automatic preservation on the blackboard, whereas visual items (e.g. the to-be-remembered information) must be subvocalized to obtain object classification.

The crucial difference between the two theories of processing lies with Jones' (1993) emphasis on the significance of seriation. He describes the incoming information as streams, which are individually programmed into linkages, connecting the objects within the concurring stream; either visual or auditory, corresponding to the order which they were perceived. The bigger the mismatch of physical properties between succeeding items within each stream, the stronger the initial linkage. Hence, this model proposes the ISE occurs due to the irrelevant sound creating competing links, which then interfere with the links of the to-be-remembered items. Thus, background noise which constitutes greater successive changes in physical properties will be more competing, subsequently causing more disruption to serial recall (Banbury, Tremblay, Macken, & Jones, 2001). According to Jones and Macken (1993), the nature of irrelevant sound should play no part in determining the extent to which serial recall is disrupted, should it be speech, tones or sound, but it is the changing state of such sound, which causes interference.

As a result of these findings, Baddeley (2007) has now discarded the argument that the irrelevant speech effect originates from phonological similarity, subsequently endorsing Page and Norris' primacy model (1998). This model proposes irrelevant sound interferes with the representation of the *order* of the to-be-remembered items, rather than with the representations of the items themselves (Norris, Baddeley, & Page, 2004; Page & Norris, 2003), much like the O-OER model. It is assumed that recall order is based on the development of associations between each successive item and the first item presented. These associations progressively decrease in strength as more items are added, offering a simple justification for memory span limitations.

Page and Norris (1998) suggested a mechanism within the phonological loop structure responsible for the orderly retrieval of items; competitive queuing, which was initially suggested by Grossberg (1978; 1987) and later advanced by Houghton (1990). The crucial element in this model is the importance of the initial item; this is because it acts as a cue for succeeding competitive queuing. The item with the

strongest association to the initial item is retrieved first; it is produced then repressed, allowing the subsequent strongest item to be retrieved. This process continues to the end of the list or to the point where the connections become too weak and the procedure collapses. Errors in recall may arise due to noise during the first stage of the process, which is responsible for order information. The association between subsequent items and the initial item may be jeopardized by noise, thus the relationship may not be as strong. Therefore according to this theory, the next item with the subsequent strongest association, which was not disrupted, is recalled instead. Unlike Baddeley's (1986) initial working memory model proposal, this adaption now provides an explanation to the issue of the irrelevant sound effect being restricted to tasks which require seriation.

An alternative theory, which may account for the ISE, is the attentional capture concept (Cowan, 1995), suggesting that a stimulus captures our attention by causing an orienting response. In relation to the ISE, it is believed that a mental template is produced by each irrelevant sound stimulus, which is then compared to each successive sound stimulus. Much like that of the changing state hypothesis, this theory proposes the change in acoustic variables in background noise, is the cause of a reduction in task performance. However, instead of an emphasis on seriation (Jones *et al*., 1992), emphasis here is laid upon attention orientation.

The theory suggests that if the mental template of a previous stimulus does not match the new stimulus presented, then orientation occurs towards the cause of the mismatch, thus the sound is attended to. Therefore, it is this orientation of attention away from the focal task, towards sound, which is argued to be the cause of the reduction in performance on concurrent tasks. However, the theory has been criticised for its emphasis on the irrelevant stimuli, consequently disregarding the necessity of seriation in the task for the ISE to be produced (Perham *et al*., 2007). Furthermore, repeated exposure to the same irrelevant stimulus, according to this theory, would predict familiarisation, thus reduction in the orientation of attention to the sound, resulting in a diminished ISE. However, it has been shown that the level of interference remains the same, regardless of exposure (Hellbrück *et al*., 1996; Jones *et al*., 1997).

### 2.5 Irrelevant Sound Effect and Speech

Among the confounding theories, there remains debate whether 'speech is special' when considering the irrelevant sound effect. The O-OER model acknowledges the greater effects which speech appear to elicit on short-term recollection, claiming speech has a higher degree of natural acoustic variability than changing non-speech items (Tremblay, Nicholls, Alford & Jones, 2000). Studies have shown, when speech and non-speech sound are made fundamentally equivalent in changing properties, through the manipulation of irrelevant sound using a signal-to-noise ratio (SNR), there are no significant differences on the effects of serial recall. Ellermeier and Hellbrück (1998) found with the addition of pink noise to a decreasing SNR, a linear improvement in recall performance was produced.

Jones *et al*. (2000) used a parametric approach, to more precisely manipulate the degree of acoustic change in speech. They used low-pass filtering, which lessened the level of distinctiveness between successive items in the to-be-ignored sound order of spoken words. In a further experiment, they used digital manipulation to

modify the degree of noise included in the sample. Increasing the irrelevant sound signal degradation revealed a linear improvement in performance for both experiments. This research endorses Jones *et al*.'s (1992) changing state hypothesis, where the main regulating feature of the interfering effects of irrelevant sound, is the extent to which the physical properties of that sound alters within the stream.

However, LeCompte, Neely and Wilson (1997) found that the use of words, particularly meaningful ones, resulted in a larger interference of recall, than tones which imitated some properties of speech. Thus it remains unclear whether there are other features of speech, besides the natural acoustic variation, which cause it to generate a greater ISE than non-speech (Little, Martin & Thomson, 2010).

## 2.6 Current Research

Many studies have compared the effects of speech and instrumental music; however few have examined the effects of vocal music on serial recall. This enables further analysis of the possible additional affects which the presence, or absence, of speech may produce.

### 2.6.1 Familiarity

Pring and Walker (1994) examined the ISE through the use of traditional instrumental music and the instrumental version of nursery rhymes, (i.e. played without the vocal component). Endorsing Baddeley's (1986) phonological loop hypothesis, they suggested that participants' memory for the lyrics of the nursery rhymes would be triggered through the unvocalised music, thus resulting in "obligatory access to phonological short-term memory" (Pring & Walker, 1994, pp. 169), consequently disrupting working memory.

Findings endorsed the predictions, where the unvocalised nursery rhymes caused more disruption than the traditional instrumental music, therefore arguing implied words, through the use of familiar music, can impair short-term recall. However, their methodology proves to have weaknesses, as there was no irrelevant speech condition for comparisons, nor was there a test for the corresponding vocalised nursery rhymes. Subsequently, their findings should be interpreted with caution, as other properties, such as differences in changing state between the traditional instrumental music and the unvocalised nursery rhymes, may be responsible for the findings, rather than the implied linguistic content. Iwanaga and Ito (2002) did, however, use both instrumental and vocal music in their short-term memory tasks. Their results showed the vocal group performed worse on verbal memory tasks when compared to the instrumental group, however, in this study, the two music conditions were of different selections, thus these findings could also have been due to a number of other physical properties instead of, or in addition to, the presence of words.

Both of the above studies exemplify issues relating to confounding variables, such as complexity and tempo. Furthermore, neither examines the effects of their sound conditions against a silent control group, in order to determine not only if, but also the degree, to which recall is disrupted by background noise. A recent study (Alley & Greene, 2008) has acknowledged the importance of these factors, improving the methodological issues identified. Instead, they examine performance in relation to

four background noise conditions: silence, vocal music, equivalent instrumental music and irrelevant speech. This enabled a comparison of the presence and absence of vocals, on otherwise equivalent instrumental music.

Similar to Pring and Walker (1994), Alley and Greene (2008) also considered familiarity with the chosen music, however they used two recent, contemporary songs, which were assumed to be "familiar to most college students" (2008, pp. 282), rather than nursery rhymes. Their aim was to determine which condition caused more interference, whether it is the vocal condition, due to complexity of the speech, or the instrumental condition, due to the 'implied lyrics'. Furthermore, it is believed that when lyrics are present, the participant does not need to search through their memory to retrieve the missing words, thus according to Alley and Greene (2008), it may be less disruptive. Their findings show the verbal condition to be more disruptive than instrumental, disputing Pring and Walker's (1994) assumption that familiar lyric retrieval interferes at a stronger degree. However, only 47% of participants reported having a high level of familiarity with the instrumental song, while 21% reported no familiarity at all. Additionally, 10% of the participants reported no familiarity with the vocal song; therefore song choice may have impacted these findings, as a proportion of participants were not familiar with the lyrics, hence would not partake in the retrieval process from long-term memory.

Although participants in Alley and Greene's (2008) study were asked to record their familiarity rating with both songs, there was no unfamiliar song for comparison of lyric knowledge. There was also no mention for consideration of the physical properties of the chosen songs, with selection based on the assumption of familiarity, thus findings may have been due to confounding variables such as complexity or tempo. Furthermore, the vocals on one song were male, and the other female, therefore this inconsistency may also have impacted on the degree of interference.

### 2.6.2 Preference

An additional characteristic within the ISE which has recently been addressed is music preference (Perham & Vizard, 2011). It has been illustrated that task performance improves when preferred music is listened to prior to task execution (Hallam, Price & Katsarou, 2002). Although, this may differ to most people's experience, as music is more likely to be listened to coincident with performing tasks, rather than prior to. The ISE, however, provides evidence that background noise causes the opposite effect on performance to occur, when played during task execution.

Perham and Vizard (2011), therefore, examined whether preference for music played during a serial recall task affected performance, compared to music which was disliked. The attentional capture theory suggests that liked music will capture attention, rather than disliked music, due to preference and familiarity (Cowan, 1995). However, findings from this research found no difference in performance between the liked and disliked music, thus Perham and Vizard (2011) concluded that preference does not affect serial recall.

The results, however, should be interpreted with care, as the manner for which the liked and disliked music was selected, was inconsistent. Participants were required to supply their own liked music, thus varying in genre, tempo and vocals, amongst other characteristics. Whereas, the disliked music was selected from a genre that

would generally not appear in the weekly music charts; thrash metal, thus was assumed to be disliked by many people. Therefore inconsistencies across the liked and disliked music may have acted as confounding variables.

In summary, former research has failed to provide corresponding evidence on the effects of vocal and non-vocal irrelevant background noise, as well as for the effects of familiarity with lyrics and songs. However, it has been identified that inconsistencies both within and between studies' methods may have influenced the confounding findings. Comparing the findings of these studies is consequently more difficult, therefore the current study served to address these weaknesses.

## 3.0 Aims

The current research aims to build upon previous studies involving the comparison of vocal and non-vocal irrelevant sound, with the refinement of some methodological issues. Unlike Pring and Walker (1994), and Iwanaga and Ito (2002), the inclusion of a silent condition in the present research will ensure the inclusion of a baseline measure of participants' serial recall, as varying levels of individual differences, in relation to working memory span, has been noted (Ellermeier & Zimmer, 1997). The current research, like Alley and Greene (2008), will examine the effects of vocal and otherwise equivalent instrumental music, while additionally examining the effects of these conditions for a familiar *and* unfamiliar song. The aim is to determine if familiarity with a song, and its lyrics, will have more or less of an effect on the serial recall of visually presented items. Thus, due to the comparison within the different sound conditions, and between the familiar and unfamiliar versions of each, the confounding findings of the effects of familiar vocal versus instrumental music (Alley & Greene, 2008; Pring & Walker, 1994; respectively), can be analysed further. While an analysis in relation to each corresponding unfamiliar condition is also enabled.

It has also been identified that copious amounts of research have examined the effects of speech and music on working memory, however as of yet, it is not apparent the two have been examined together, as one condition. This research, therefore, has incorporated additional conditions, where lyrics of both the familiar and unfamiliar songs have been recorded as speech. Furthermore, the speech has been recorded over the instrumental music of the corresponding song, replacing the sung vocals. This will enable comparisons of familiarity with the lyrics, while also providing a novel condition of speech over instrumental music, to examine whether speech may be special, with possible additional interference. Hence working memory performance will be examined under six conditions: silence, white noise, original, equivalent instrumental music, recorded spoken lyrics and spoken lyrics over instrumental music.

Unlike Alley and Greene (2008), songs were selected with regard to consistency of otherwise confounding variables, as well as familiarity levels, thus vocals for both songs were female, and consequently recorded lyrics were also female. In addition, taking Jones *et al*.'s (1992) changing state hypothesis into consideration, where the amount of disruption to serial recall is related to the degree and amount of change between beats, songs were chosen on the number of beats per minute (bpm), making them as similar as possible.

In relation to preference, the inconsistencies with song choice identified in Perham and Vizard's (2011) study have been accounted for in this research. The inclusion of a questionnaire in the current research allows the level of preference to be determined, with the intention of distinguishing whether preference for background music affects performance on serial recall tasks.

This research, therefore, primarily predicts (i) all sound conditions will have a significant, detrimental effect on recall compared to silence, while, (ii) the addition of spoken lyrics over the instrumental music will produce the largest damaging effect on recall. It is also predicted that (iii) exposure to any of the conditions within the familiar song, will result in a poorer performance when compared to its corresponding unfamiliar condition. Finally, (iv) preference for background music will result in a higher degree of interference.

## 4.0 Methods

A pilot study was executed prior to the subsequent research to check and alter detected methodological issues.

### 4.1 Participants

Twenty-four undergraduate students (11 Males; mean age 23: minimum 20, maximum 36; standard deviation 3.61), from various universities across Edinburgh, took part in the experiment. The participants were recruited through opportunistic sampling over a 2 week period in February 2012. Each participant reported normal or corrected-to-normal vision and hearing.

### 4.2 Apparatus and Materials

#### 4.2.1 To-be-remembered material

The working memory task presented participants with a short-term word recall task similar to that used by Marsh *et al.* (2009). The word sets were generated using the MRC Psycholinguistic Database (Wilson, 1988), which is a machine usable dictionary. Criteria of minimum 5 and maximum 6 letters per word, to reduce confounding variables in relation to the word length effect was selected, along with a familiarity rating of 400-600 (printed familiarity); concreteness rating of 500-600 and imageability rating of 500-600 (ratings derived from merging Pavio, Yuille & Madigan, 1968; Toglia & Battig, 1978; and Gilhooly & Logie, 1980 norms). As the maximum possible rating for each was 700, these ratio-style scores were selected on the basis of gaining a substantial amount of narrowly ranging, highly familiar, concrete and imageable words.

Ninety-six words from the produced list were then selected at random, constructing 12 word lists, of 8 words each. [The pilot study comprised word lists of both 12 and 10 items, in an attempt to prevent ceiling effects regarding Miller's (1967) magic number 7 +/- 2, however both proved too long for participants' short-term memory, with only recency effects being produced. Therefore word lists were further reduced to 8 for the continuation of the research.] Word lists were randomly constructed, then checked for obvious semantic relations (e.g. *train* and *ticket* were randomly assigned in the same word list, and so were separated to different lists). They were then

pseudo-randomly arranged, so that no adjacently presented words were of phonological similarity. Colours as a category of high frequency words, was used for the practice trial composing 5 words.

### 4.2.2 Irrelevant sound

All participants were presented with silence and white noise generated from *Audacity*: a sound editing program (available from: http://audacity.sourceforge.net/). The remaining sound conditions were constructed from two songs; "We Found Love" by Rihanna ft. Calvin Harris and "Sans Souci" by The Red Velvetines. The former was the 5$^{th}$ biggest selling single of 2011 in the UK (The Official UK Charts Company, 2012), and so was used with the assumption the majority of university students would be familiar with it. The latter is a single from an unsigned band from Northern Ireland, and was used with the assumption that participants would not have heard it before and so would be unfamiliar with the track.

Both songs comprised four sound conditions; the original version, an (otherwise) equivalent instrumental version, the spoken lyrics and the spoken lyrics over instrumental. The songs were of similar bpm, where the familiar song was set with 128bpm, and the unfamiliar was 134bpm. Both songs had female vocalists, and so the recording of the spoken lyrics was also female, furthermore a Scottish female, as the research was executed in a Scottish university this is the accent most subjects would encounter on a daily basis. The spoken versions were recorded also using the *Audacity* software. Each piece of audio lasted 27 seconds and comprised the same section for each song. The audio was played using Labtec Spin 95 PC multimedia speakers.

### 4.2.3 Questionnaire

Participants were required to complete a short questionnaire after the experiment. They were asked seven questions to determine their familiarity with the two songs, along with their music preferences (Appendix D). They were questioned if they could identify either song, and then asked to record their familiarity with the songs and their knowledge of the songs' lyrics separately on a 5-point Likert scale to verify that they were/were not familiar with the associated song.

## 4.3 Design

The experiment had a mixed-subjects design, whereby each participant was subject to all of the possible six sound conditions; silence, white noise, original, instrumental, spoken lyrics and speech over instrumental. However, they were differentiated depending on whether they heard the familiar or unfamiliar version for each sound condition (see Appendix E). Two of the twelve to-be-remembered lists were randomly assigned to one of the six sound conditions. The silent condition was executed first, as a baseline measure for short-term serial recall, followed by white noise. The order of the remaining four sound conditions was counterbalanced across all participants, to control for potential order effects. As a compromise for repeated measures, the four sound conditions relative to the familiar and unfamiliar songs were split, so that each participant heard all four conditions; 2 were of the familiar song and the remaining 2 of the unfamiliar song, also counterbalanced across all participants.

### 4.4 Procedure

Participants provided informed consent to partake in the research, with the knowledge that withdrawal was possible at any time (Appendix A & B). They were seated at a viewing distance of approximately 60cm from a *PC* monitor, where the to-be-remembered stimuli would appear in a central position. Word lists were presented in upper case black 72-point *Arial* font, one word at a time, against a white background. A 3 second delay occurred once the experiment began, before the appearance of the first word, to prevent any affects which may arise due to distraction, or surprise, when a new auditory extract began. Each word appeared on screen for 2 seconds, with an inter-stimulus interval of 1 second. Recollection was immediate, with the end of the list being alerted by the presentation of a red 'RECALL', similar to that of Marsh *et al*. (2009) methods. Participants had 15 seconds to recall in the practice, and 30 seconds for the remaining trials.

Participants were tested individually, in a silent, private computer lab, to reduce the effects of confounding variables. Participants were informed they would be presented with 12 word lists of 8 words, one word displayed at a time. They are asked to recall these in serial order, on the sheets provided, once the 'RECALL' cue was displayed on screen. Recall sheets comprised 12 columns of 8 rows each, and 1 of 5 rows for the practice. Participants were asked to inform the experimenter when they were ready to begin each trial. No audio was played for the practice trial however when applicable, it manually began coincident with the start of the trials, ceasing when the red 'RECALL' cue was presented, lasting 27 seconds. Participants were informed they would have 30 seconds to recall as many of the words as possible, in the order and placement in which they were presented. Participants were not informed of the conditions they would be exposed to, however, were instructed to ignore any background audio which may be played throughout the entirety of the experiment, in order to avoid participants consciously focusing on what is intended to be irrelevant background noise. Once participants completed the recall trials, they were asked to fill out a short questionnaire, and were provided with a debrief sheet on exit (Appendix C). The pilot study highlighted issues with the length of the word lists only, therefore the remaining procedure resumed unchanged. The experiment lasted approximately 25 minutes.

## 5.0 Results

Recall measures came in two forms; the number of successive correctly recalled items ("serial"), and the number of correctly recalled items in its corresponding place, regardless of missing items ("placement"). Scores from the two trials for each condition were then totalled.

### 5.1 Descriptive Statistics

On initial exploration the data were found to be normally distributed, with the removal of outliers; participant 24 in regards to their placement scores over the sound conditions, and their instrumental score when exploring familiarity, along with the speech over instrumental placement score for participant 19 in regards to familiarity analysis.

Table 1 below shows the means and standard deviations of each of the six conditions, including the separate scores in relation to familiarity for each of the song conditions. The silent condition shows the highest mean recall for both the serial and placement scores, while the lowest came from the speech over instrumental condition, showing a larger difference between the two serial mean recall scores at 5.08, while placement mean recall difference is 3.63. Placement mean recall scores were higher than serial across all conditions, while both scores show lower recall for each of the four versions of the familiar song compared to the unfamiliar song, suggesting the familiar song caused more disruption than its corresponding unfamiliar conditions. The mean placement scores for the original and instrumental versions of the unfamiliar song were both higher than those of the white noise condition, 10.17, 9.55 and 9.5 respectively. This pattern, however, was not evident for serial recall, with the scores progressively getting lower from white noise to the original and finally the instrumental conditions, which is also evident for the familiar song. The largest difference between the unfamiliar and familiar songs occurred in the instrumental version, with a serial score difference of 2.25, suggesting the familiar instrumental version caused more interference than its counterpart unfamiliar condition.

**Table 1**
**Mean and standard deviation scores for serial and placement recall in each sound condition**

| | | Serial Score | | Placement Score | |
|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* |
| **Sound Condition** | | | | | |
| **Silent** | | 9.12 | 1.30 | 10.67 | 1.81 |
| **White Noise** | | 7.92 | 1.84 | 9.50 | 1.62 |
| **Original** | Familiar | 5.83 | 2.66 | 8.42 | 2.61 |
| | Unfamiliar | 7.08 | 2.54 | 10.17 | 2.25 |
| | **Total** | 6.46 | 2.62 | 9.29 | 2.55 |
| **Instrumental** | Familiar | 4.42 | 2.47 | 8.33 | 1.44 |
| | Unfamiliar | 6.67 | 2.43 | 9.55 | 1.92 |
| | **Total** | 5.54 | 2.65 | 8.91 | 1.76 |
| **Spoken** | Familiar | 4.17 | 2.41 | 6.83 | 3.04 |
| | Unfamiliar | 5.17 | 2.04 | 7.50 | 1.78 |
| | **Total** | 4.67 | 2.24 | 7.17 | 2.46 |
| **Speech over Instrumental** | Familiar | 3.42 | 2.61 | 6.83 | 2.86 |
| | Unfamiliar | 4.67 | 2.71 | 7.27 | 2.49 |
| | **Total** | 4.04 | 2.68 | 7.04 | 2.64 |

### 5.2 One-way ANOVA

A one-way repeated measures analysis of variance (ANOVA) was carried out to establish whether there were significant differences between the six sound conditions, initially irrespective of familiarity.

#### 5.2.1 Serial Recall

Mauchly's test of sphericity (Appendix F) showed serial recall scores over the six sound conditions to be significant with $p$ <.05, therefore the null hypothesis of homogeneity of covariance is rejected and sphericity is not assumed, thus Greenhouse-Geisser is used instead. The ANOVA showed a significant main effect of sound condition on the serial recall of items, F(3.78, 86.83) = 20.66, $p$ <.01, $\eta_p^2$ = .47, a large effect size (Kinner & Gray, 2011), with post hoc tests (Bonferroni) revealing significant differences (displayed in Table 2). Significant differences were found between the silent condition and all other sound conditions, while the speech over instrumental condition was also found to be significantly different to white noise ($p$ <.01) and the original ($p$ <.05) conditions, however not against the instrumental or spoken alone.

#### 5.2.2 Placement Recall

When analysing the placement scores in relation to the sound conditions, Mauchly's test of sphericity (Appendix F) was not significant with $p$ >.05, therefore the null hypothesis of homogeneity of covariance is accepted and sphericity is assumed. A significant main effect of sound condition on the placement recall of items was found, F(5, 110) = 14.99, $p$ <.01, $\eta_p^2$ =.41, a large effect size (Kinner & Gray, 2011). Table 2 shows the significant differences for placement scores, repeating the serial score results for the silent condition, excluding a significant difference between silent and original conditions where $p$ =.29. A significant difference between the speech over instrumental condition and both white noise and original conditions were found, however there was also trend found between the speech over instrumental and instrumental conditions, along with a trend for instrumental and spoken conditions.

**Table 2**
**Bonferroni's Post-hoc results for comparisons between sound conditions for both serial and placement scores**

| Sound Condition | Score | Silent | White Noise | Original | Instrumental | Spoken |
|---|---|---|---|---|---|---|
| **White Noise** | Serial | .014 | | | | |
| | Placement | .03 | | | | |
| **Original** | Serial | .00 | .283 | | | |
| | Placement | .291 | 1.00 | | | |
| **Instrumental** | Serial | .00 | .023 | 1.00 | | |
| | Placement | .006 | 1.00 | 1.00 | | |
| **Spoken** | Serial | .00 | .00 | .112 | 1.00 | |
| | Placement | .00 | .002 | .079 | .054 | |
| **Speech over Instrumental** | Serial | .00 | .00 | .013 | 1.00 | 1.00 |
| | Placement | .00 | .00 | .014 | .056 | 1.00 |

### 5.3 Familiarity

When considering the mean scores for both serial and placement measures in relation to familiarity and sound condition, Figures 1.1 and 1.2 below clearly shows that participants continuously performed better when under the unfamiliar condition, than for the familiar, for both serial and placement recall. In relation to participant's preference for listening to either the unfamiliar or familiar song, a significant difference of $p$ <.01 was found, with preference for the familiar song (Appendix F). Thus although participants perform worse under the familiar conditions, they prefer to listen to that music.

According to participants self-report of familiarity with the unfamiliar song, all reported a score of 1; indicating they were not familiar with the song or the songs' lyrics. In relation to the familiar song, 79% of participants gave a score of 5, indicating they were very familiar with the song, while 17% gave a score of 4 and one participant reported a score of 3, indicating a moderate familiarity. In relation to knowledge of the lyrics, nearly half (46%) reported knowing all of the lyrics, a third reported a score of 4 while 21% reported partial knowledge of lyrics with scores of 3 or 2.

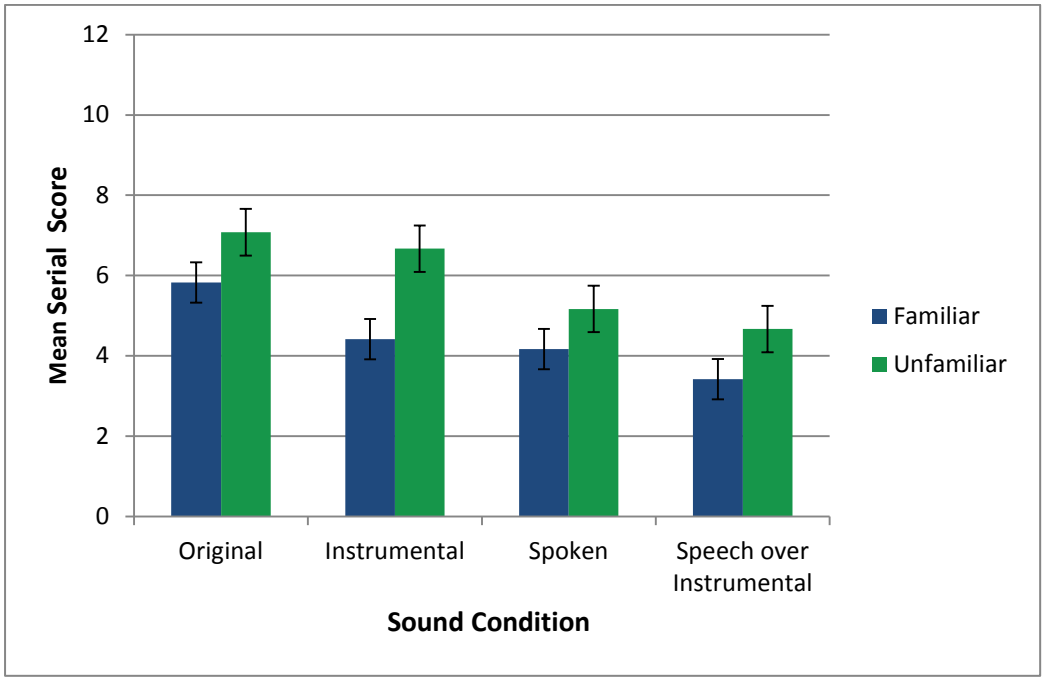


**Figure 1.1: Graph of familiar and unfamiliar mean serial score for the corresponding four sound conditions**

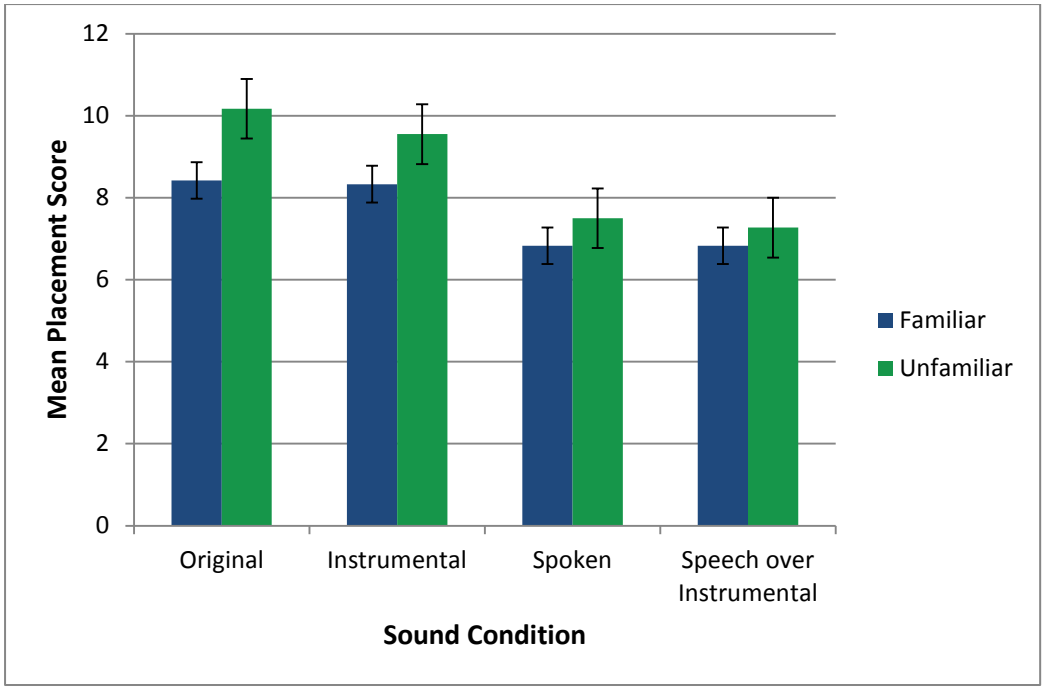


**Figure 1.2: Graph of familiar and unfamiliar mean placement score for the corresponding four sound conditions**

### 5.4 Two-way ANOVA

A two-way between subjects ANOVA was executed in order to determine whether there was an interaction between the four sound conditions and familiarity for the serial and placement recall scores.

#### 5.4.1 Serial Recall

A significant main effect for the serial scores relating to sound was found, $F(3, 88) = 4.29$, $p < .01$, $\eta_p^2 = .13$, a medium effect size (Kinner & Gray, 2011), and for familiarity, $F(1, 88) = 8.01$, $p < .01$, $\eta_p^2 = .08$, a medium effect size (Kinner & Gray, 2011), however no significant effect was found for an interaction between the two, where $p = .83$.

#### 5.4.2 Placement Recall

Analysis of the placement scores identified a main significant effect for sound, $F(3, 86) = 5.77$, $p < .01$, $\eta_p^2 = .17$, a large effect size (Kinner & Gray, 2011), and for familiarity, $F(1, 86) = 4.36$, $p < .05$, $\eta_p^2 = .48$, a large effect size (Kinner & Gray, 2011), however there was also no significant effect found for the interaction between the two, where $p = .78$.

Comparisons between the sound conditions for both serial and placement scores are displayed in Table 3, showing significant differences. The original song and spoken lyrics over the instrumental showed a significant difference for both serial and placement scores, whereas placement recall scores produced further significant differences between original and spoken conditions, and the instrumental and spoken lyrics over the instrumental condition.

**Table 3**
**Bonferroni's Post-hoc results for significant condition comparisons**

| Comparisons between sound conditions | Significance |
|---|---|
| **Serial** | |
| Original – Speech over Instrumental | $p < .01$ |
| **Placement** | |
| Original – Speech | $p < .05$ |
| Original – Speech over Instrumental | $p < .01$ |
| Instrumental – Speech over Instrumental | $p < .05$ |

### 5.5 Multivariate ANOVA

Finally, a multivariate ANOVA was carried out to determine whether there were significant differences within each condition in relation to familiarity. The only condition to display a significant difference between familiar and unfamiliar serial scores was instrumental at $p < .05$ (Appendix F).

## 6.0 Discussion

It has previously been identified that serial recall is susceptible to interference from irrelevant auditory stimuli, with the degree of hindrance dependant on the variation in acoustic properties of the sound (Baddeley 1986; Jones & Macken, 1993; Jones *et al*., 1993). One aim of this investigation was to distinguish whether familiarity with a song, and its lyrics, had an effect on serial recall. Furthermore, to examine whether the effect was more or less disruptive when compared to a corresponding unfamiliar condition. Another objective of this research was to explore the degree to which familiar and unfamiliar speech was detrimental, through the inclusion of the spoken lyrics over instrumental music and spoken lyrics alone conditions.

As expected, participants were found to perform best under the silent condition. In comparison, the spoken lyrics over instrumental condition elicited the worst serial recall performance, supporting predictions. Additionally, the mean performances for familiar sound conditions were consistently lower than those of their corresponding unfamiliar condition, also verifying predictions. Finally, participants reported a preference for the familiar rather than unfamiliar song. Therefore, when considered alongside the consistently poorer performance for familiar conditions, this finding indicates that exposure to preferred background music results in a more detrimental effect, than exposure to disliked music, confirming the final hypothesis.

### 6.1 Silence is best

When examining the mean scores for both the serial and placement recall measurements, it is clear that performance is best when there is no background noise present. This is in line with previous research on the irrelevant sound effect (Alley & Greene, 2008; Baddeley 1986; Jones & Macken, 1993), indicating that silence is key for optimal short-term serial recall. This could be due to a number of

reasons, however it is only through examining the effects of differing background sounds, in relation to performance, that it may be understood why.

The mean scores also identified white noise to successively produce the least disruption to performance, after silence. This supports the changing state hypothesis (Jones *et al*., 1992), where it is argued that differing acoustic properties, within a sound stream, cause the ISE. Furthermore, it is the extent to which succeeding tones differ from one another, that determines the extent to which serial recall is disrupted. Hence, in relation to white noise; which consists of a fixed bandwidth at a centre frequency, (i.e. minimal fluctuation), less disruption was observed when compared to the other sound conditions, however, significantly more interference was observed when compared to silence. Further endorsing that silence offers the best environment for working memory performance.

### 6.2 Importance of Speech

The spoken lyrics over instrumental condition was found to produce the lowest mean recall scores for performance, further supporting the changing state hypothesis. According to this theory, the altering physical properties in irrelevant sound is the *sole* determinant of the detrimental effects on serial recall (Jones *et al*., 1992). Hence, due to the addition of the natural acoustic complexity of speech (Tremblay *et al*., 2000), over the fluctuating tones in the instrumental music, this particular condition could be argued to possess the largest overall changing state, consequently producing the poorest recall.

However, this finding may be due to a number of other reasons. As Baddeley (1986) had originally proposed, some form of auditory filter may be regulating the sounds which gain access to his proposal of the phonological loop. It was believed that speech, or speech-like sounds, gain automatic access to the phonological loop (Larsen *et al*., 2000). Therefore in relation to this finding, it may be a case that speech; due to automatic access, along with the addition of the seriation of instrumental music, (more recently identified by Baddeley (2007) as a further disruptive force of serial recall), may have produced the poorest mean scores.

An additional explanation for speech over instrumental music causing the largest degree of disruption may be due to the infrequent exposure participants would experience to such a condition in everyday life. Participants would not be used to the lyrics of songs being recorded in speech, rather than sung vocals, over the instrumental music, thus this specific condition may lack ecological validity. However, speech may occur *on top* of original versions of songs on a daily basis, e.g. in the workplace. Hence, to examine the impact of speech alongside vocal music, future research should include such a condition to facilitate comparisons.

Further highlighting the impact of speech on the ISE, performance in the spoken lyrics condition was found to produce the successive poorest mean recall scores. When comparing the two features individually, which caused the poorest mean recall when combined; speech, and instrumental music, the former caused a higher degree of interference. Furthermore, a trend for the placement score is suggested for the difference in performance between the two conditions (see Table 2). These findings suggest that speech, either due to the natural fluctuation, or possibly automatic access, is more detrimental that just tones.

However, the original versions of both songs produced the best performance in comparison to the other song conditions. Also, a significant difference was found between the original version of the song, and the spoken lyrics over instrumental condition (see Table 2). The changing state hypothesis argues that acoustic fluctuation causes such results; however, the vocals in the original version also display changes in state, in addition to the changing tones. Thus, this significant difference may not be explained through the changing state hypothesis, but instead may be more closely related to Baddeley's (1986) auditory filter hypothesis. These findings may suggest that it is not the linguistic aspect of such material which causes extensive disruption, but that specifically *speech* may be 'special', in relation to the ISE.

Considering an evolutionary view, we may be more tuned to permit speech automatically to mental working space, due to its use as a method of communication; therefore, more of the limited capacity space in working memory may be utilized. However, previous findings have displayed the lack of irrelevant sound effect that one speech utterance produces (Jones & Macken, 1993; Jones, *et al*., 1992; Jones, *et al*., 1993). Although, it could be suggested that the repetition of a single word, may insinuate that the unattended processing of that word, along with recognition from long-term memory, may occur just once. Following processing, the word may be interpreted in relation to its acoustic characteristics; where no changing state would occur, until the auditory properties change, hence explaining the lack of detrimental effect found in previous research. The mental space, therefore, would not be utilized to the same degree as processing not only various words, but also the semantic connection between them.

In relation to this, future research should examine the effects of speech utterances further, possibly through alternating between two words initially, followed by lengthening the list. Bearing in mind the changing state hypothesis, phonological complexity should be taken into consideration to reduce confounding variables, thus single syllable words may be best to test for processing, rather than supplementary alterations in the acoustic properties.

However, previous research has shown semantic and phonological similarity between the to-be-remembered and to-be-ignored material produces no additional interference on performance (Bridges & Jones, 1996; Jones & Macken, 1995a; LeCompte & Shaibe, 1997). Although, analysis of the effects of semantic and phonological similarity *within* the to-be-ignored material, and its effect on serial recall, has not been thoroughly addressed within the literature. Therefore, with the addition of words to the lists of speech utterances previously suggested, analysis of nonsense sentences as irrelevant background material, compared to semantically related sentences, would enable examination determining whether the relationship of words within the irrelevant background noise, has a differing effect on serial recall.

When considering previous research concerning speech, which employed signal-to-noise ratio (SNR) methods (Ellermeier & Hellbrück, 1998; Jones *et al*., 2000), no significant differences in performance were found for speech and non-speech sound. This method entails manipulating the properties of speech to essentially possess the same as non-speech sounds. It can be argued, through the lessening of the fundamental physical characteristics of speech, the natural features which make it distinct from other sounds, are removed. Thus, if speech is special, in regards to

automatic processing, the features which make it so may have been erased. This method essentially reduces speech to general sound, therefore the perception of speech as a form of communication would be jeopardised and so would be processed more like tones than words.

The current research is from a behavioural perspective, however, when considering evidence for the processing of speech from a neuropsychological perspective, functional imaging research has suggested that higher order linguistic information is processed differently to that of low-level acoustic information (Davis & Johnsrude, 2003). Thus, indicating that speech and non-speech sound are processed differently. Furthermore, it is suggested that two separate functional routes exist; one for the lexical-semantic processing of speech, and one for the auditory-motor association (Ward, 2006). The former route is believed to make contact with structures which are crucial for semantic memory, such as the anterior temporal lobe (Scott & Wise, 2004). Whereas, it is within the latter route, that Hickok and Poeppel (2004) have suggested a neuroanatomical basis for the phonological loop; between the parietal and frontal lobes. Therefore this suggests the involuntary processing of sounds may occur along the auditory-motor association route, whereas the processing of speech would occur along the lexical-semantic route.

These suggestions from neuropsychological research support the suggestions made in the current study, where it is believed that speech may be 'special' relative to processing, compared to other auditory stimulus. However, care should be taken when interpreting the findings of Davis and Johnsrude's (2003) study, as they examined the effects of *attended* background noise, rather than unattended. Therefore examination of the effects of speech and non-speech as irrelevant stimuli should be executed to determine whether these findings remain constant for unattended material.

### 6.3 Commentary on Familiarity

This study also examined the effects of familiarity with songs and their lyrics on serial recall performance. Participants expressed high levels of familiarity with the familiar song and lyrics, while all participants indicated they were not familiar with the song, or lyrics, of the chosen unfamiliar song. Therefore familiarity levels of both songs are reliable, suggesting comparisons of the recall scores for the familiar and unfamiliar conditions have face validity.

When comparing performance in relation to familiarity, participants who were exposed to the unfamiliar conditions, consistently out preformed those in the corresponding familiar condition (see Figures 1.1 & 1.2). This suggests familiarity with background noise is more disruptive for serial recall tasks, than that of unfamiliar sound. Furthermore, participants reported a preference for the familiar song over the unfamiliar. This is in line with the attentional capture concept, where it is believed that liked music will cause orientation of attention towards it, to a higher degree, than that of disliked stimuli, due to familiarity and preference (Cowan, 1995). Performance in the focal task is therefore threatened, resulting in a decline. However, this theory has been criticised for placing too much emphasis on the irrelevant stimulus, subsequently neglecting the importance of seriation in the task, which has been displayed to be crucial for the effect to occur (Perham *et al*., 2007). Therefore, this explanation may lack specificity to the ISE.

Previous research, regarding preference for background music, found no significant difference in serial recall performance between liked and disliked music (Perham & Vizard, 2011). However, the auditory stimulus used was inconsistent for the familiar condition; therefore it is difficult to compare the results to these findings in relation to the current study. It should be noted, however, the preference results in the current study may have merely been due to the music being familiar. Therefore auditory stimuli which contain a range in both familiarity, and preference, should be employed for future research concerning this domain.

While many confounding variables were accounted for in the current research, such as consistency with female vocals, and similar beats per minute for the chosen songs, not all acoustic characteristics could be controlled. Therefore other properties, such as frequency, may have contributed to the familiarity finding. However, the same pattern was found for the spoken lyrics condition, which excluded any instrumental stimuli; therefore familiarity could still justify these findings.

When considering the changing state hypothesis, familiarity with background noise should not affect performance on serial recall, as it is exclusively the fluctuation in the physical properties which do so, due to the conflict of two seriation streams (Jones & Macken, 1995a). Thus, due to consideration for the controlled variables listed above, the level of disruption for each corresponding condition should have been similar; however this was not the case. Furthermore, the instrumental condition produced a significant difference of performance between the familiar and unfamiliar condition. Therefore, the changing state hypothesis cannot account for this finding.

In relation to previous research regarding the effects of familiar sound, findings from this study exhibit a parallel finding with Pring and Walker (1994), where the familiar instrumental condition was more disruptive than the original. Although, caution should be taken in regards to comparing current results with those form Pring and Walker's (1994) study, as the instrumental conditions were not compared to silence, or to the corresponding vocal version of the nursery rhyme. The current findings dispute that of Iwanaga and Ito's (2002), and Alley and Greene's (2008) results, where the opposite was true. It is suggest their findings support Baddeley's working memory model, where the presence of the linguistic aspect of the song, rather than implied familiar lyrics, is principally responsible for the degeneration in performance.

However, as identified, within the song conditions of the current research, the original version exhibited the least decline in performance, relative to silence. Pring and Walker (1994) argue that the implication of words that familiar instrumental music generates consequently triggers the missing lyrics to be retrieved from long-term memory, resulting in obligatory access to the short-term phonological store. This, in turn, requires additional processing from the limited capacity working memory, hence less mental working space is available for the rehearsal of the to-be-remembered material, subsequently producing a poorer recall performance. Thus, Baddeley's working memory hypothesis still holds true for these findings, however, is interpreted in a different way. However, it should be noted that there may be different levels of familiarity with the stimulus used in Pring and Walker's (1994) study; nursery rhymes, and that of the current; contemporary chart single. Nursery rhymes are typically learnt in childhood, thus have been stored in memory for much longer

than that of the more recent familiar song used in the current study. Comparisons are then limited due to the differing familiarity levels.

Interestingly though, the unfamiliar instrumental condition also provoked more disruption than the unfamiliar original condition. This finding cannot be justified by Pring and Walker's (1994) implication of familiar words, as the lyrics are unknown to the participants. Nor can the changing state hypothesis provide explanation, as according to this theory, the original version should include more changing states than the instrumental alone, regardless of familiarity, thus should elicit a poorer performance.

Although it is uncertain why the unfamiliar sound follows the same degenerative pattern as familiar sound, a similar process which occurs for the implied lyrics hypothesis, may provide insight. Just like the searching of long-term memory for familiar lyrics, the participants, in this case, may be searching their memory to recognise the song, by searching for lyrics which may fit the tune. However, due to unfamiliarity, no words in long-term memory are a successful match, thus no additional words would enter the phonological loop.

This suggestion may also provide an explanation for why a significant difference between the two instrumental groups was discovered. According to this, lyrics for the familiar condition are accessed and subsequently enter the phonological loop, whereas no words are identified for the unfamiliar condition, therefore none enter the phonological loop. Searching long-term memory utilizes cognitive processing; however, the addition of successfully retrieved words for the familiar condition, produces more pressure on the limited capacity cognitive space. Consequently, this could explain why the familiar instrumental condition produces a further decline in recall, compared to the unfamiliar instrumental condition.

### 6.4 Limitations

The main limitation of the study relates to the recording of the spoken lyrics. A fellow Psychology student, from Edinburgh Napier University, recorded both sets of spoken lyrics. Many of the participants, due to opportunistic sampling, were also Psychology students from the same year. Consequently, recognition of a fellow student's voice may have been more off-putting, or oriented participants' attention towards the irrelevant stimuli, to a higher degree. Thus, the findings related to speech are questionable, as confounding variables, such as familiarity with the voice used, may have accounted for the high levels of disruption. Future research should therefore use a voice which would not be familiar with possible participants.

Furthermore, irrelevant background noise was manually started for relevant conditions. Although great care was taken to start the sound stimulus coincident to the start of each trial, and a 3 second delay between the start of a trial and the presentation of the 1$^{st}$ word was enforced, it was impossible for timing to be accurately consistent throughout the experiment. However, this was due to the choice of software, as it was not possible to combine the audio material with the visually presented words. Thus use of a computer programme which enables the combination, would ensure consistency for future research.

It has also been identified that the location from where the disruptive auditory stimulus is orientated, affects the degree of recall. Buchner, Bell, Rothermund and

Wentura (2008) found a worse impairment on serial recall when the irrelevant sound was directionally close to where the to-be-remembered items were visually presented, compared to when the audio came from behind. As mentioned, the audio for the current research began manually, and the direction from which was behind where participants executed the task. This finding suggests that the level of disruption observed in the current research may be understated, thus future research should take care in the orientation of irrelevant auditory stimulus.

There were a total of 6 conditions, which consisted of 2 trials each, thus participants were required to partake in 12 recall tasks, within a short time period of approximately 15 minutes. Although all conditions and trials were counterbalanced to control for potential order bias, participants' performance may have improved with experience, with possible development of strategies. Therefore results may be an artefact of this experience. It is suggested for future research to reduce the number of trials required of participants, with possibly the addition of a further silent condition at the end of the experiment, to compare with baseline scores for experience effects.

Furthermore, although significant differences were found, the sample size was relatively small at 24. Also, due to opportunistic sampling, all participants were university students, thus the sample may not be representative of the whole population and so may not be generalizable. It is therefore suggested for future research to include a wider range of participants to obtain more valid results.

As noted earlier, there are individual differences in relation to the ISE (Ellermeier & Zimmer, 1997). One such difference involves the method which participants utilized to retain the to-be-remembered material. There is some evidence in the current findings which suggests participants did not use rehearsal as a retention method, but semantic chaining instead. When analysing answers on the recall sheets, some semantic errors were made, for example participant 6 recalled *jelly* followed by *fish* rather than the correct item: *fiddle,* while participant 7 recalled *monster* instead of *beast*. Interestingly, two of the participants recalled *seller* rather than *cellar* displaying the correct recall phonologically, thus suggesting covert phonological rehearsal was executed. As stated in the methodology, care was taken to prevent semantic relation within the word lists; however, it appears individuals use semantic chaining, however obscure, as a method of retention. Such differences are difficult to control, due to the personal, internal nature of them, however, instruction to restrain from doing so, before taking part in the task, may aid in the reduction of such a retention method.

### 6.5 Future Research

Future research could take one of two suggested routes. As mentioned above, in relation to further analysing the effects of speech on the ISE, speech utterances should be re-analysed. The alteration between two words, with the addition of words to the list, and consideration for phonological complexity, mean examination for the involuntary semantic processing of words, in relation the ISE, would be enabled. Furthermore, also adopting a neuropsychological research method, such as brain imaging, would permit analysis for identifying possible anatomical locations for the involuntary processing of speech and non-speech during working memory tasks.

An alternative route which future research could adopt, relates to preference and familiarity to background stimuli. Confounding findings within the literature suggests

this area requires more analysis. Consideration for the degree of familiarity of the irrelevant sound should be adopted. As identified, Pring and Walker's (1994) study used nursery rhymes as familiar instrumental music. As most individuals are introduced to these as children, the lyrics have been stored in memory for much longer then the familiar audio used in the current study. Although current results endorsed those of Pring and Walker's (1994), comparisons are difficult due to the difference in the level of familiarity. Therefore future research should employ a range in familiarity levels of stimuli; a contemporary single, controlled exposure to an unfamiliar song for an intermediate condition, along with a never-heard-before unfamiliar song. The inclusion of a nursery rhyme would also enable comparisons to previous research, while enabling different levels of familiarity to be assessed. Furthermore, preference for irrelevant stimuli should also be analysed, separate to that of familiarity, as to avoid familiarity acting as a confounding variable to the preference measure.

Analysis of these different elements which cause the ISE separately, and a comprehensive understanding of the extent to which these factors affect serial recall, will allow for a more inclusive examination of the factors together in the future.

### 6.6 Implications of study

Findings identified in the current study may have implications for study skills which students may adopt for examination preparation. If a student's revision process relies on remembering information in serial order, such as the sequence to a mathematical equation, then playing music in the background will disrupt the process. What's more, if the songs which are played are familiar, this research suggests this will have an additionally detrimental effect, thus, if music is desired while studying, then listening to unfamiliar music would threaten performance to a lesser extent.

Furthermore, employee performance in the workplace may also be jeopardised by background music. For workplaces which usually expose their employees to music, e.g. retail, the retention of information, such as item codes, would be vulnerable to interference. Not only are such employees exposed to music, but also speech, thus the findings from this study would indicate a higher level of disruption to working memory would occur. Furthermore, due to the ISE currently being a laboratory based phenomenon, more research should be executed in order to determine the detrimental effects of involuntary processing of background music in everyday situations, such as in the workplace.

Additionally, as previously identified, seriation is inextricably related to language learning and production (e.g. Gupta & MacWhinney, 1997; Jones, *et al*., 2006; Perham, *et al*., 2009; Schweppe & Rummer, 2007), thus playing music while a child learns their first language may be detrimental.

## 7.0 Final Conclusion

Although the sample size was small in the current study and findings need to be replicated in order for definitive conclusions to be attained, the present report offers evidence which favours the view that speech is 'special' and may play a critical role in the disruptive effects of irrelevant sound. Furthermore, it appears that familiar

stimuli elicit a higher degree of interference, than that of unfamiliar stimuli, in relation to serial recall.

Current findings therefore suggest it is not solely the degree of fluctuation within the unattended background sound which affects serial recall, but that familiarity and speech may have special, yet somewhat undefined roles, in the disruption of working memory. Extensive analysis of the particular effects of familiarity and irrelevant speech is suggested. Specifically, the effects of semantically related compared to nonsense sentences as irrelevant stimuli. Furthermore, a multidisciplinary approach, including behavioural and neuropsychological methods, would enable a more comprehensive understanding of the precise effect speech and non-speech material elicits on serial recall.

## References

Alley, T. R., & Greene, M. E. (2008). The relative and perceived impact of irrelevant speech, vocal music and non-vocal music on working memory. *Current Psychology, 27,* 277-289.

Baddeley, A.D. (1986). *Working memory.* Oxford: Clarendon Press.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417-423.

Baddeley, A.D. (2003). Working memory: Looking back and looking forward. *Nature Reviews, Neuroscience, 4*, 829-839.

Baddeley, A.D. (2007). *Working memory, thought, and action.* Oxford: Oxford University Press.

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.H Bower (Ed.), *The psychology of learning and motivation, Vol 8.* London: Academic Press.

Baddeley, A.D., & Salamé, P. (1986). The unattended speech effect: Perception or memory? *Journal of Experimental Psychology*, *12*, 525–529.

Beaman, C.P., & Jones, D.M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state speech hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 459–471.

Beaman, C.P., & Jones, D M. (1998). Irrelevant sound disrupts order information in free as in serial recall. *Quarterly Journal of Experimental Psychology, 51A*, 615–636.

Buchner, A., Bell, R., Rothermund, K. & Wentura, D. (2008). Sound source location modulates the irrelevant sound effect. *Memory & Cognition*, *36*, 617-628.

Banbury, S., & Berry, D. (1998). Disruption of office- related tasks by speech and office noise. *British Journal of Psychology, 893,* 499-517.

Banbury, S.P., Tremblay, S., Macken, W.J., & Jones, D.M., (2001). Auditory distraction and short-term memory: phenomena and practical implications. *Human Factors, 43*, 12–29.

Bridges, A.M., & Jones, D.M. (1996). Word-dose in the disruption of serial recall by irrelevant speech: Phonological confusions or changing state? *Quarterly Journal of Experimental Psychology, 49A,* 919-939

Colle, H.A., & Welsh, A. (1976). Acoustic making in primary memory. *Journal of Verbal Learning & Verbal Behavior*, *15*, 17-32.

Cowan, N. (1995). *Attention and memory: An integrated framework.* New York: Oxford University Press.

Davis, M., & Johnsrude, I.S., (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience, 23*(8)*,* 3423–3431.

Ellermeier, W., Hellbrück, J., (1998). Is level irrelevant in "irrelevant speech"? Effects of loudness, signal-to-noise ratio, and binaural unmasking*. Journal of Experimental Psychology: Human Perception and Performance, 24*(5)*,* 1406–1414.

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the "irrelevant speech effect." *Journal of the Acoustical Society of America, 102*, 2191-2199.

Gilhooly, K.J. and Logie, R.H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words*. Behaviour Research Methods and Instrumenta- tion, 12*, 395-427.

Grossberg, S. (1978). Behavioral contrast in short-term memory: serial binary memory models or parallel continuous memory models? *Journal of Mathematical Psychology, 3,* 199-219.

Grossberg, S. (1987). Competitive learning from interactive activation to adaptive resonance. *Cognitive Science, 11,* 23-63.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267–333.

Hallam, S., Price, J., & Katsarou, G. (2002). The effects of background music on primary school pupils' task performance. *Educational Studies, 28*, 111–122.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language *Cognition, 92*, 67-99.

Hellbrück, J., Namba, S., & Kuwano, S. (1996). Irrelevant background speech and human performance: Is there long-term habituation? *Journal of the Acoustical Society of Japan*, *17*, 239-247.

Houghton, G., (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish & M. Zock (Eds.) *Current research in natural language generation.* London: Academic Press.

Iwanaga, M., & Ito, T. (2002). Disturbance effect of music on processing of verbal and spatial memories. *Perceptual Motor Skills, 94*, 1251–1258.

Jones, D.M., Alford, D., Bridges, A., Tremblay, S., & Macken, W.J. (1999). Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 464-473.

Jones, D.M., Alford, D., Macken, W.J., Banbury, S.P., & Tremblay, S., (2000). Interference from degraded auditory stimuli: linear effects of changing-state in the irrelevant sequence. *Journal of the Acoustical Society of America, 108 (1*), 1082–1088.

Jones, D.M., Hughes, R.W., & Macken, W.J. (2006). Perceptual organization masquerading as phonological storage. *Journal of Memory and Language, 54*, 265–2281.

Jones, D.M., & Macken, W.J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory*. Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 369-381.

Jones, D.M., & Macken, W.J. (1995a). Phonological similarity in the irrelevant speech effect: Within-or between-stream similarity*. Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 103–115.

Jones, D.M., Macken, W.J., (1995b). Organizational factors in the effect of irrelevant speech: the role of spatial location and timing. *Memory & Cognition, 23*, 192–200.

Jones, D.M., Macken, W.J., & Mosdell, N. (1997). The role of habituation in the disruption of recall by irrelevant sound. *British Journal of Psychology*, *88*, 549-564.

Jones, D.M., Macken, W.J., Murray, A.C., (1993). Disruption of visual short-term memory by changing-state auditory stimuli: the role of segmentation. *Memory & Cognition, 21*, 318–328.

Jones, D.M., Madden, C., Miles, C., (1992). Privileged access by irrelevant speech to short-term memory: the role of changing state. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 44A*, 645–669.

Kinnear, P.R., & Gray, C. D. (2011). *IBM SPSS statistics 18 made simple.* Hove: Psychology Press.

Larsen, J.D., Baddeley, A.D., & Andrade, J. (2000). Phonological similarity and the irrelevant speech effect: Implications for models of short-term verbal memory. *Memory, 8*, 145-157.

Le Compte, D.C., &Shaibe, D.M. (1997). On the irrelevance of phonological similarity to the irrelevant speech effect. *Quarterly Journal of Experimental Psychology*, *50A*, 100–118.

LeCompte, D.C., Neely, C.B., & Wilson, J.R. (1997). Irrelevant speech and irrelevant tones: The relative importance of speech to the irrelevant speech effect*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 472–483.

Little, J.S, Martin, F.H, & Thomson, R.H.S. (2010). Speech versus non-speech as irrelevant sound: Controlling acoustic variation. *Biological Psychology, 85,* 62–70.

Logie, R.H. (1996). The Seven Ages of Working Memory. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus & R.T. Zacks (eds.), *Working memory and human cognition.* Oxford University Press: New York

Miyake, A. (2001). Individual difference in working memory: Introduction to the special selection. *Journal of experimental Psychology (General), 130*(2), 163-168.

Norris, D.G., Baddeley, A.D., & Page, M.P.A. (2004). Retroactive effects of irrelevant speech on serial recall from short-term memory*. Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1093–1105.

Page, M., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, *105*, 761–781.

Page, M.P.A., & Norris, D.G. (2003). The irrelevant sound effect: What needs modelling, and a tentative model. *Quarterly Journal of Experimental Psychology*, *56A*, 1289–1300.

Pavio, A., Yuille, J.C. and Madigan, S.A. (1968). Concrete-ness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement, 76* (3, part 2).

Perham, N., Banbury, S.P., & Jones, D.M. (2007). Reduction in auditory distraction by retrieval strategy. *Memory, 15*, 465–473.

Perham, N., Marsh, J.E., & Jones, D.M. (2009). Syntax and serial recall: How language supports memory for order. *Quarterly Journal of Experimental Psychology, 62*, 1285–1293.

Perham, N. & Vizard, J. (2011). Can preference for background music mediate the irrelevant sound effect? *Applied Cognitive Psychology, 25*, 625–631.

Pring, L., & Walker, J. (1994). The effects of unvocalized music on short-term memory. *Current Psychology*, *13*, 165–171.

Salamé, P. (1990). Effects of music, speech-like noise, and irrelevant speech on immediate memory. In B. Bergland & T. Lindvall (eds.), *Noise as a public health problem*. Stockholm: Swedish Council for Building Research.

Salamé, P., & Baddeley, A.D. (1982). Disruption of short-term memory by irrelevant speech: Implications for the structure of working memory. *Journal of Verbal Learning & Verbal Behavior*, *21*, 150-164.

Salamé, P., & Baddeley, A.D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society*, *24*, 263-265.

Salamé, P., & Baddeley, A.D. (1989). Effects of background music on phonological short-term memory. *Quarterly Journal of Experimental Psychology*, *41A*, 107–122.

Schweppe, J., & Rummer, R. (2007). Shared representations in language processing and verbal short-term memory: The case of grammatical gender. *Journal of Memory and Language, 56,* 336–356.

Scott, S.K., & Wise, R.J. (2004). The functional neuroanatomy of prelexical processing in speech perception. *Cognition, 92*, 13-45.

The Official UK Charts Company (2012). Available from: http://www.theofficialcharts.com/chart-news/the-top-20-biggest-selling-singles-of-2011-revealed/ Accessed on: 15.01.12

Toglia, M.P. and Battig, W.R. (1978). Handbook of Semantic Word Norms. New York: Erlbaum.

Tremblay, S., Nicholls, A.P., Alford, D., Jones, D.M., (2000). The irrelevant sound effect: does speech play a special role*? Journal of Experimental Psychology: Learning, Memory & Cognition, 26,* 1750–1754.

Ward, J. (2006). *The student's guide to cognitive neuroscience*. Hove: Psychology Press.

Wilson, M.D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioural Research Methods, Instruments and Computers*, *20*(1), 6-11.