Exploring a Process-Based Account of the Disruption to Music Cognition by Task-Irrelevant Sound

by

Rona Doris Linklater

A thesis submitted in partial fulfilment for the requirements for the degree of Doctor of Philosophy at the University of Central Lancashire.

May 2020

STUDENT DECLARATION FORM

Type of Award **Doctor of Philosophy**

School **Psychology**

1. Concurrent registration for two or more academic awards

I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution.

2. Material submitted for another award

I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work.

3. Collaboration

Where a candidate's research programme is part of a collaborative project, the thesis must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

NOT APPLICABLE

4. Use of a Proof-reader

No proof-reading service was used in the compilation of this thesis.

Signature of Candidate

Print name: Rona Doris Linklater

Abstract

Music is ubiquitous and important in everyday life and yet a complete understanding of the role it plays in cognitive processes and distraction remains elusive. This thesis aimed to explore the extent to which background melody and lyrics—alone or combined within song—impact differentially on concurrent cognitive processes. Current theoretical accounts question specificity for music and language by arguing that lexical, phonological and music processing share a common cerebral network: yet other lines of evidence indicate that separate working memory processes for music and visual-verbal information exist. Numerous studies involving to-be-ignored sound have considered its auditory properties and the interplay between rehearsal, attention, or task complexity. However, most prior research addressing interference produced by music on task performance has focused on short-term memory recall/recognition of visually presented tones/words. Few studies address vocal production of melody/lyrics and consequently it is still unclear how pathways for vocal input/output are generated/related and, of greater consequence, how the vocal-motor planning mechanism required for vocal production is affected by the competing motor-plan from the presence of extraneous sound: melody, song, or speech. To explore distraction within a larger process-based cognitive framework, the present empirical series of studies are the first to demonstrate effects of to-be-ignored distracters on long-term memory retrieval and production for complete melody and lyrics of known songs through vocal (humming) and speaking performance. Various combinations of to-be-ignored familiar and unfamiliar melody, with and without lyrics, were delivered during retrieval performance of a different familiar target melody (Study I) and lyrics (Study II). Results suggest an independence of language and melody processing and are consistent with an interference-by-process framework: the nature of the disruptive effects observed on the focal task were jointly dependent on the nature of the distracters and the focal task. However, using short-term memory tasks, Studies III and IV provide some evidence against the interference-by-process view. The results extend the perceptual-gestural view of short-term memory, according to which the disruption observed by task-irrelevant sound reflects a clash between the action of the sequencing processes embodied within perceptual input-processing and gestural output-planning systems that are general and co-opted to meet task demands. Implications and applications of these results are discussed with regard to age-related cognitive decline.

Table of Contents

List of Tables

List of Figures

This thesis is dedicated to my wonderful mum Doris Linklater (1912 – 2012) whose smile will be with me forever

And to my dad William Ronald Linklater (1911 – 1988) to make him proud

"He who hears Music, feels his solitude peopled at once" - *Robert Browning.*

Acknowledgements

I would like to record heartfelt gratitude to Dr John Everett Marsh, my Director of Studies, for the invaluable guidance and encouragement shown to me throughout this PhD course of study. Without your patient understanding, meticulous attention to detail, and thoughtful constructive feedback, this journey would not have been possible. You saved my life and took me down a pathway hitherto unknown but totally challenging and rewarding. Your calm approach, stanch support, and unquestionable dedication to academic excellence was inspirational and provided the motivation to learn and improve. “Tack så mycket John”. I also wish to acknowledge the wise words and encouragement from Dr Jeannie Judge. Your insight and reassuring support was greatly valued—particularly your critical eye and helpful suggestions during the last stages of Chapter writing.

I owe my survival over the last six years to several good friends who became actively engaged in the research. Grateful thanks to Carol Rose, for your expertise in providing the vocals on all the auditory excerpts—which was definitely a fun experience for both of us. To Anne-Louise Coldicott for your unwavering encouragement as both friend and musical assessor, and to Cath, Janet, Liz, Ruth, and Sylvia, for your invaluable work as assessment raters and participants. Finally, thank you Tanya for understanding my moments of anxiety on the UCLan journey, and Shaun for your kind supportive words. However, none of this would have been possible without assistance from the army of willing participants, whose generosity of time, cheerfulness throughout the “ordeal”, and positivity towards the research aims was humbling. Your contribution provided for a quality of research data that is most valued and appreciated.

CHAPTER I
INTRODUCTION

Music and cognition

This thesis aims to explore the role of different perceptual and cognitive processes in the remembrance of song. In order to advance understanding of how retrieval processes relating to lexical/lyrical/melodic content operate and interact, patterns of retrieval in the presence of potentially distracting task-irrelevant sound will be examined. Examining the propensity for distraction via differing properties of task-irrelevant sound will allow some evaluation of whether perceptual and cognitive systems are shared, or overlap for music and language processing, or are discrete and non-overlapping. For this thesis, distraction is the key theoretical tool of measure, whereby distraction via differing properties of task-irrelevant music and language, presented during long-term semantic memory focal retrieval task performance (Study I-II), or during presentation of visual items for short-term immediate recall (Study III-IV), should enable a better understanding of the role of perceptual or vocal-motor processes such as rehearsal, attentional processes (e.g., attentional capture) and/or their interaction upon mnemonic pathways for music/language. Each of four empirical studies will address potential variation in distraction effects produced by manipulations of familiar and unfamiliar melody, sung-lyrics, and spoken-lyrics. Collectively, these studies compare memory recall for both serial and non-serial tasks and consider the disruption to these tasks in light of existing theories and mechanisms of distraction. Throughout the four studies it is expected that the interplay of the vocal-motor production mechanism for familiar song and its vulnerability to disruption via the presence of familiar and unfamiliar task-irrelevant music/language will be exposed.

Chapter I comprises an outline of ways in which the brain responds to auditory disruption through music and language retrieval. This will be discussed first within the music domain and will be evaluated based on perception, processing, and production requirements of the two dominant features of music, pitch, and rhythm. Following this exposition there will be coverage of the purported shared cognitive resources of music and language within song. An examination of semantic memory processes will lead to those aspects pertaining specifically to engagement of the motor systems by identifying links between auditory perception, semantic selection, and vocal-motor production (Tsang, Friendly, & Trainor, 2011). Chapter I concludes with an overview of relevant aspects of music support to aid language retention in cognitive decline when word production may have become impaired.

The geneses of music and language, and how the brain responds, is still fuelling debate (McDermott & Hauser, 2005; Mithen, 2005; Molino, 2000; Tomasello, Velichkovskiĭ, & Rumbaugh, 1996; Wallin, Merker, & Brown, 2000). Nevertheless, it is apparent that language, music or non-music, has the potential to influence and possibly control aspects of human thinking (Whorf, 1956). Song combines music—melody, and language—lyrics. Both independent fundamentals are themselves hierarchical systems that have comparable properties (Schön, Magne, & Besson, 2004). For example, discrete items (notes of pitch, words) are combined into larger structures such as musical chords, phrases, or sentences (Fiveash, McArthur, & Thompson, 2018). Rules, acknowledged as syntax, govern how this combination operates (Fiveash et al., 2018; Thaut, 2009). Evidence to date regarding their representation within memory, is not resultant from vocal production outcomes in terms of performance accuracy from long-term memory, but is largely based on either violation (e.g., Fiveash et al., 2018; Miranda & Ullman, 2007; Perruchet & Poulin-Charronnat, 2013), or musical/verbal interference within short-term memory recognition studies (e.g., Thompson & Yankeelov, 2012). These studies have not provided unequivocal support for the notion that music and language are stored/represented separately or together within memory. Also, the idea of a single mechanism for music does not take account of its complexity (pitch, duration, expression, emotion). A more appropriate explanation would be the presence of brain specialization, a "modular architecture", for local neural circuitries essential for music processing (Peretz, 2003; Peretz & Coltheart, 2003, p. 689).

Therefore, to bridge this chasm in the literature, the following series of empirical studies sought to investigate, for the first time, both short-term memory recall tasks (Study III-IV), and long-term memory for music and language combined within familiar song (Study I-II) requiring retrieval and production of complete familiar melodies and associated lyrics. The intention of this undertaking was to identify similarity and difference of retrieval patterns in relation to the presence of task-irrelevant melody, sung-lyrics, and spoken-lyrics. Further to perception and recognition of familiar target items an added component for these empirical studies is vocal recall performance. Therefore, the reported findings will include analysis of the interaction between perception and action. Musical performance is complex and cognitively challenging (more so in the presence of a different irrelevant music-related performance) with processing that requires intricate timing of hierarchical organizations and auditory motor-interaction (Pfordresher, 2006; Zatorre, Chen, & Pehune, 2007).

1.1 Neural response to music

The interaction, and subsequent organisation of sounds into an ordered structure, categorizes sounds as music. “Music can be thought of as a type of perceptual illusion in which our brain imposes structure and order on a sequence of sounds” (Levitin, 2006, p.109). Rhythm (duration) and pitch (tone) are the two primary musical elements and a mass of research into their processing has been conducted via neurological study. Some argue that memory for rhythm and pitch are treated as one unit (Jones & Boltz, 1989), however consensus would seem to indicate some independence (Krumhansl, 2000; Peretz, 1990; Peretz & Coltheart, 2003). Time relations, or rhythm, has two main features. A steady regular meter (containing a combination of strong and weak beats), and duration of individual note-lengths (time values, grouped into rhythmic patterns). Separation of meter and duration identified through findings that showed tapping a rhythmic pattern with the right hand and the beat with the left hand was easier than vice-versa (Ibbotson & Morton, 1981). Hence, Ibbotson and Morton (1981) deduced the right hemisphere was engaged for metre and the left hemisphere for rhythm. This was reinforced by Sakai et al. (1999) who, via neuroimaging, identified different frontal and cerebellar mechanisms were engaged for rhythms that were metric or nonmetric.

Although pitch processing is paramount to music processing, it is intervallic relationships between pitches that define melodic contour (or diatonic/chromatic chords of harmony) to determine familiarity (Hébert & Peretz, 1997). It would appear when reliant on contour representations the right superior temporal gyrus is crucial in melody recognition, but when reliant on interval information, without contour cues or scalic information (Tillman, Bharucha, & Bigand, 2000), right and left temporal areas are activated (Peretz, 1990; Peretz, 2001) and involved in recognition. A lack of awareness of scale structure, therefore, can result in an inability to distinguish tonal from atonal movement or retain memory for pitch (Besson & Faïta, 1995; Peretz, 1993). This has led to the suggestion, following functional Magnetic Resonance Imaging studies (fMRI[1]; Janata et al., 2002), that there are separate neural networks dedicated to processing melodic scale structure. Following an experiment with brain damaged patients using positive emission tomography (PET[2]) scans, Peretz (1990) identified segregation between pitch arrangements to form a melody and duration (rhythm). This double dissociation between melodic and temporal processing led Peretz (1990) to

---

[1] fMRI is a procedure that measures brain activity by detecting changes in blood flow in response to neural activity.

[2] PET is an imagery test that detects changes in chemical activity to determine neural firing.

deduce distinct neural mechanisms specific to music processing. Brain damage which interfered with pitch discrimination spared aspects of time relationships and conversely, rhythmic components may become impaired while pitch content is spared (Midorikawa, Kawamura, & Kezuka, 2003; Peretz, 1990).

Much discussion regarding deficits in music and language processing is reported from neuropsychological studies of patients with amusia (e.g., Peretz, 2003). Amusia has become a generic term, corresponding to musical aphasia, covering a complex group of impairments to musical ability for perception, processing, and production (Benton, 1977). This term embraces rhythm, pitch, familiar tune recognition, and an ability to sing/hum, resulting from right temporal lobe lesions (Henson, 1977). One instance was noted in the French composer Ravel, who demonstrated aphasia for processing compositional rule (Alajouanine, 1948, cited in Nunes-Silva & Haase, 2013). If people with brain damage can recognize melodies presented without words, that were previously very familiar to them, but can no longer recognize spoken-lyrics, voices, and other sounds, it would appear that distinct processing modules do indeed exist for music and speech (Schlaug, Jancke, Huang, & Steinmetz, 1995). Identification of a double dissociation between vocal production for speaking and singing support this concept (Peretz et al., 2009). Following evidence that people with particular brain damage can recognise a melody when accompanied by associated lyrics, but not when lyrics are withheld, Nunes-Silva and Hasse (2013) argue "musical lyrics are processed in parallel to musical melodies" (p. 49) but, according to the model suggested by Peretz and Coltheart (2003), in a different language processing system[3].

The way in which music is pro-active in stimulating varied human responses is dependent on those brain areas triggered by the musical stimuli, not just the auditory cortex (Patel, 2008). For example, music can connect to the nucleus accumbens, associated with aesthetic rewards, seeming to identify listening to music as an enjoyable and rewarding experience (Blood & Zatorre, 2001; Salimpoor & Zatorre, 2013). It also has potential to impact one's emotions, evoke memories, and influence one's well-being (Levitin, 2006; Sachs, 2006, 2007). Using fMRI, Alluri (2011) revealed how areas responsible for emotions, creativity, and motor actions are activated when the brain processes music elements such as rhythm, tonality, or timbre. Whilst individuals listened to a modern Tango, Alluri (2011) identified music "activates not only the auditory areas of the brain but also employs large scale neural networks" (cited in

[3] Modular music processing model see Chapters III and IV.

Science Daily, 2011, p. 1). Moreover, brain responses of individuals listening to musical genres from Vivaldi to the Beatles showed instrumental and vocal music to be processed differently (Alluri et al., 2013), lyrics shifting processing of musical features towards the right auditory cortex. Similar neural mechanisms to those linked with pleasant/unpleasant emotions also appear to be engaged by music (Blood, Zatorre, Bermudez, & Evans, 1999). Identification of processing disparities produced by hearing task-irrelevant music with/without spoken or sung-lyrics and their potential to arouse associated memories is a fundamental facet of this thesis.

1.2 Memory processes

Much early experimental and clinical research into brain activity shows a left hemisphere specialization for language/speech with right hemisphere specialization for other non-linguistic operations (Bever & Chiarello, 2009). For the majority of people, the left hemisphere, in particular Broca's area, associated with processing grammar, and Wernicke's area, associated with vocabulary, deals with the more academic activities of language such as reasoning, logic, number, and analysis (Rathus, 2012). Images, imagination, recognition, and the perception of music require the right hemisphere (Buzan, 1983). However, more recently activation in both Broca and Wernicke's areas has been shown for some aspects of music processing (Koelsch et al., 2002; Levitin & Menon, 2003). These findings question domain specificity (Schön et al., 2010).

Music is a dual hemispheric skill. Although music shares language properties with the left hemisphere, such as temporal order/rhythm, interaction occurs with the right hemisphere, important for perception of non-verbal dynamic, timbre/emotion (Gates & Bradshaw, 1977; Gordon, 1981; Segalowitz, 1983). Gates and Bradshaw (1977) identified the neuromuscular aspects of singing to be controlled by the right hemisphere following findings that a patient was able to sing familiar songs despite left hemisphere removal. In addition, right hemisphere domination has been shown, for people without musical training, when required to memorize unfamiliar melodies, not based on semantic strategy (Zattorre, Evans, & Meyer, 1994), and for unknown melody recognition by people with either-hemispheric brain damage (Samson & Zatorre, 1991; Zatorre, 1985). It appears that when judging music familiarity, the left hemisphere appears to be more influential (Platel, 2005). Furthermore, in a two-tone probe identification task, a typically left hemisphere related interval-based route was used by musicians (Bever & Chiarello, 1974). Although neither hemisphere appears dominant for music processing (David, 1994) there stands a consensus, derived from studies utilizing fMRI, that right-lateralized temporal, frontal and parietal brain areas are most significant for pitch contour recognition (Peretz, 2003). However, if interval information is required, both hemispheres are implicated (Ayotte, Peretz, Rousseau, Bard, & Bojanowski, 2000). In sum, “music perception supports the hypothesis that the left hemisphere is dominant for analytic processing and the right hemisphere for holistic processing” (Bever & Chiarello, 2009, p. 94).

1.2.1 Working memory and long-term memory.

The Working Memory model (WM; Baddeley & Hitch, 1974) is a dominant account of short-term memory that pervades current theorising concerning the processing of phonological and spatial material in immediate memory. This multi-component approach to short-term memory identifies a phonological loop (comprising an articulatory control process and phonological store), as one component of two storage systems (along with the visuo-spatial sketchpad “inner eye” responsible for processing of visual and spatial material) which, together with the central executive control system (that deals with cognitive demands) form the model. Components comprise a “passive” phonological store, to which speech is assumed to have direct access but visually presented information requires phonological recoding (grapheme to phoneme) to be granted access, an “inner ear” dealing with speech like sounds, and an articulatory loop or “inner voice” which “rehearses” sound for up to two seconds (Baddeley & Hitch, 1974). This “inner voice”, or articulatory control process, as part of the phonological loop is consequently linked to speech production (Smith, Wilson, & Reisberg, 1995). This sub-vocal process of WM explains some physical movements during a task, such as a slight movement of the larynx, detected by electromyography recordings (EMG; Locke & Fehr, 1972). The articulatory process is responsible for internal speech, performed without sound, which operates during a task such as reading a book or following a music score (Cleland, Davies, & Davies, 1963; Rayner & Pollatsek, 1989). Silent or internal speech may also contribute to access of meaning and understanding (Rayner & Pollatsek, 1994). In the context of short-term memory, the articulatory process revivifies decaying memory traces that are within the “passive” phonological store. Without rehearsal, the items are lost altogether from the store after 2-seconds.

The proposed partnership between the inner voice and the inner ear and the notion of direct access of auditory material *versus* indirect access of visual material, can explain a range of short-term memory phenomenon (Baddeley, 1986; but see Jones, Macken, & Nicholls, 2004) including, for example, the articulatory suppression effect.

This refers to the finding that recall of letters, digits, or musical sequences are impaired by constantly repeating a task-irrelevant utterance such as "the" during the encoding of to-be-remembered items (Murray, 1967; Saito, 1997, 1998; Salamé & Baddeley, 1989; Schendel & Palmer, 2007). On the WM model, this articulatory suppression effect occurs because repetition of a task-irrelevant word occupies the articulatory process thereby preventing the entry of verbal material into the phonological store (Baddeley, 2000). Further, the WM model explains why short-term memory is poorer for phonologically similar as compared to dissimilar lists of items (the phonological similarity effect). This is due to a confusion between the phonemes of similar item representations within the phonological store (Baddeley, 1986). The WM also explains why the phonological similarity effect is reduced under conditions of articulatory suppression with visual presentation (since the items of visual origin cannot enter the phonological store) and interact with auditory presentation (since auditory items gain direct access to the phonological store wherein they can become confused; for an alternative account see Jones et al., 2004).

The partnership between the inner voice and the inner ear can also explain a number of so-called secondary phenomena. These include, for example, why preventing subvocalization impairs the parsing of meaningful letter strings (e.g., deciphering NRG to "energy") and scanning familiar melodies (e.g., generating a melody from a song-title to judge whether a given note falls or rises compared with its predecessor). In this process, participants sub-vocally generate or rehearse imaged material (using the inner voice; articulatory control process), permitting a judgement to be made by the inner ear (the phonological store; Smith et al., 1995).

The WM model is pertinent to the upcoming studies and will be dealt with in greater detail in Chapter II-III. However, suffice to say, for now, is that there is good evidence that sensorimotor mechanisms become activated by a relevant verbal-visual presented title (just as with auditory imagery generated from a song-title in order to judge rising/falling pitch; Smith et al., 1995) and while hearing task-irrelevant familiar music (with or without associated lyrics). In the latter case, exposure to task-irrelevant familiar music may drive its perception and trigger internal production of melody/lyrics ("inner singing"; Dalla-Bella, Tremblay-Champoux, Berkowska, & Peretz, 2012). A consequence of this, as will be drawn out in Chapter II-III, is that an ensuing competition would arise between irrelevant melody/lyrics and a target material with the likely consequence of impairing or delaying production of target material (e.g., melody/lyrics).

Over the last twenty years, WM theorists have proposed the existence of additional storage possibilities within the memory system. For example, Baddeley (2000) proposed that visual and phonological information could be integrated into a coherent whole, or multi-modal code, within an "episodic buffer" that acts as an interface for the other slave systems (e.g., phonological loop, visuo-spatial sketchpad) and long-term memory (LTM). The role of the episodic buffer is, therefore, to communicate between LTM and other WM components, and to store bound visual and phonological memories into an "episode" or story (Baddeley, Allen, & Hitch, 2010).

In this approach, it is possible that the episodic buffer is active in supporting an unfolding melody/lyrics sequence during retrieval/performance, from LTM (as required for Study I-II in this empirical series) as compared to during an imitation task when short-term memory (STM) is engaged in temporary, immediate repetition of to-be-recalled sequences (required for Study III). However, the episodic buffer view has been questioned since it predicts a close relationship between executive function and LTM that is not always observed (Berlingeri et al., 2008; Gooding, Isaac, & Mayes, 2005). For example, Berlingeri et al. (2008) found a group of people with Alzheimer's disease (AD) to be proficient in tasks requiring central executive functions but showed no immediate prose recall, generally thought to require some LTM ability.

In relation to memory for non-verbal musical stimuli, some music psychologists have argued for WM independence of musical from that of verbal material (Berz, 1995). Rather than having one acoustic store for language and music, Berz proposed a separate component, an additional slave system: a music memory loop comprising a music store and control process within the WM model. He suggested chunks of memorized musical information, a hierarchical schema, organized and held within LTM, influence encoding of music in STM. This would then account for irrelevant instrumental music having a different disruptive effect on verbal performance, as opposed to irrelevant vocal music or irrelevant speech (see Chapter II for relevant distraction studies). Although some research suggests that musical sounds share the same processing stages as non-musical sounds (Peretz & Zatorre, 2005), there are some aspects of musical/verbal segregation owing to unique processing required of musical pitch stimuli (Peretz, 2003), that suggest language and melody are processed independently (Besson, Faita, Perez, Bonnel, & Requin, 1998; Bonnel, Faita, Peretz, & Besson, 2001; Racette & Peretz, 2007; Roberts, 1986; Schulze, Muller, & Koelsch, 2011).

1.2.2 Memory for music and language.

Unlike spoken language, music is not one identifiable cognitive capacity but is multi-dimensional, comprising pulse, pitch, duration, phrasing, tempo, and timbre together with an emotional component. The combination of these has resulted in an assorted system of rules, the “grammar” of music (Thaut, 2009). When we hear music a rapid creation of a temporally organised pattern of sounds must be assembled with perception triggering cognitive analysis. Human ability to develop these into musical language results from an inherent ability to “think music”. Thaut (2009) compares this development with our verbal and number language systems. Therefore, if music is thought of as a core language, its influence as “precursors and prerequisites for the development of higher cognitive executive functions” could be possible (Blacking, 1973; Cross & Woodruff, 2009). Konecni (1982) states that cognitive capacity is required for processing music and, resulting from its multi-dimensional composition, music memory embraces all aspects of LTM. Semantic memory is necessary for pitch contour, rhythmic shape, and phrase structure, episodic memory for prior performance and associated autobiographical details, and implicit procedural memory for requisite performance skills without awareness of conscious thought (Cohen & Squire, 1980).

1.2.3 Similarities and differences.

Similarities between language and music domains may be identified through their rule-governed hierarchy. Both are based on rules that combine into higher-order structures according to rules of syntax (e.g., notes and chords into musical phrases, phonemes and words into sentences). Although some syntactical structures are universal, others are cultural (Chomsky, 1988). In addition, they both depend on long term memorised representations, such as d-o-g is a four-legged creature, and a familiar melody is not just series of pitch, but a combination of pitch, rhythm, and form, evolving over time in memory as a progression in serial order. Both disciplines have a temporal aspect, whereby a series of pitch, or phoneme sequences, unfold through time, and contain an element of emotion or prosody (Besson & Schön, 2001).

However, there are fundamental differences. For example, the metric structure through a section of music is fairly consistent but is less specific in language. Each musical component comprised within a musical phrase (e.g., rhythm, pitch, harmony) is heard in combination and each component may carry its own syntax (Besson & Schön, 2001). Rules of musical grammar are therefore, arguably, more prone to ambiguity than syntactic rules of language (Aiello, 1994). Music also contains a vertical dimension that is not present in language. Dissimilar notes or words, within a vocal counterpoint, such as the polyphonic madrigal writing of Morley, may be pleasing to the ear, unlike the cacophony created by dissimilar words concurrently produced by different speakers. Although, both music and language carry expectation for what is to follow (Meyer, 1956), music conveys meaning through memory and emotion as, unlike speech, it does not possess a semantic system. However, one area of consensus is that music and language cannot be considered as discrete entities (Besson & Schön, 2001). Temporal, melodic, and harmonic aspects of music may each involve a different method of processing. Likewise, language processing needs to consider subdivision of phonemes into morphemes into words and then into meaningful sentences. Furthermore, the syntax of music, in relation to chord formation, harmonic progression, or rhythmic and pitch patterning, has been shown to interact with processing of language syntactic information (Koelsch, 2011).

1.2.4 Specificity of music and language processes.

Syntactic processing.

Much evidence for specificity of music and language processing has resulted from neurological research using violation tasks and brain imagery. Event-related brain potential (ERP[4]) studies have shown similar brain responses to structural elements of language (e.g., syntax) and music (e.g., harmony) during tasks containing violations to verb/noun order or "wrong", non-diatonic, chord notes (Besson & Faita, 1995; Janata, 1995; Kutas & Hillyard, 1983). These results contrast with outcomes from brain responses to written or spoken semantic incongruities (Kutas & Hillyard, 1980).

The shared syntactic integration resource hypothesis (Patel, 2008), suggests domain specific representational networks that contain information specific to either music or language, and general resource networks that enable syntactic integration into sequences dependent on the auditory grouping of elements into a single auditory stream (Bregman, 1990; Koelsch, 2011). An alternative theory, the syntactic equivalence hypothesis (Koelsch, 2013), supports shared syntactic processing but also suggests that these are not shared by semantic or acoustic deviance processing. Both theories have been supported through syntactic violation experiments (e.g., Fiveash & Pammer, 2014; Kunert, Willems, Casasanto, Patel, & Hagoort, 2015), and neuroimaging studies (e.g., Koelsch, Gunter, Wittfoth, & Sammler, 2005b; Schön, et al., 2004). However, Kunert et al. (2015) in a fMRI study demonstrated that when simple and complex sentences were sung to melodies in key, or with one note out-of-key, or with increased dynamic, the complex sentences with an out-of-key note produced an overlap in the Broca's area

[4] ERP measures the time course and scalp distribution of the brain's response resulting directly from a specific stimulus.

that did not occur for the dynamic condition, thereby demonstrating a syntactic-specific interaction.

When processing structure of rule-based knowledge (Seidenberg, 1997), similar cognitive operations are involved whereby syntactic violations to language and music produce similar brain responses (Janata, 1995; Kutas & Hillyard, 1983). For example, Miranda and Ullman (2007) found a double dissociation from neuroimaging studies between knowledge of rule and memory, previously found in language (Newman, Pancheva, Ozawa, Neville, & Ullman, 2001), was also prevalent in music following familiar/unfamiliar melody in and out of key note-violation tasks. Their results indicated a double dissociation for rule governed and memory-based melodic processing analogous to that shown for rule and memory within grammar and lexical processing.

Encoding of pitch involves some melodic features common to music and language, such as pitch contour and intervallic relationship used in speech prosody. In a neurological listening study, that required participants to determine correct final notes/words, Schön et al. (2004) compared prosody and melody, wherein emotional intonation contours of speech and melodic phrases identified variations in pitch, amplitude, and duration, for musicians and non-musicians. Pitch manipulation was either congruous, weakly incongruous (35% - 1/5 of a tone), or strongly incongruous (120% - a semitone, e.g. F# replacing F). Electrophysiological data identified the pitch processing reaction time. A parametric F0 manipulation revealed similar cognitive processes engaged in violation of pitch in both domains suggesting shared neural resources, with musicians more accurate at recognising pitch violation in language and music. In addition, only musicians recorded faster reaction times for the stronger incongruities. These results contrast to findings of impaired melodic but intact prosodic pitch processing (e.g., Peretz et al., 2002), although task difficulty may have been influential here, as musicians can easily detect a quarter tone pitch change whereas for a question, intonation rise is nearer to an octave. Nonetheless, processing semantic aspects of language and music during violation tasks have shown dissimilar results. Violations to grammatical rules have produced brain area reactions that did not occur for violations to memory-based knowledge from lesion studies supporting some distinction in the neural bases for language and music processing (Newman et al., 2001; Ullman et al., 1997).

Semantic processing.

In neurological studies, an N400[5] component is established as a measure of the integration process for language processing (Kutas & Hillyard, 1980). The N400 occurs when semantically unexpected/incongruous words are present within a linguistic context (Besson & Faïta, 1995). Whilst a N400 can be generated from ERPs by semantically incongruous words (e.g., "He takes coffee with cream and *sugar/dog*", Kutas & Hillyard, 1980), incongruities within music (e.g., a non-diatonic note) are normally associated with a P300[6] (Besson et al., 1998). Violation studies using ERPs have identified some evidence for specificity of semantic music and language processing (Besson & Schön, 2001). Besson and Schön (2001, Experiment 1) sought to investigate whether the N400 established for language processing (Kutas & Hillyard, 1980) could be replicated when melodic and harmonically unexpected notes occurred within a familiar and unfamiliar melodic phrase (Besson & Faita, 1995). There were three end-note conditions: expected note, in-key unexpected, non-diatonic unexpected. P600s[7] were identified for both unexpected phrase endings, the level dependent on the degree of incongruity. The late positive component (LPC) had a larger amplitude and a shorter latency to the non-diatonic than diatonic incongruity. The larger amplitude and shorter latency of the LPC for incongruous notes ending familiar musical phrases, as compared to unfamiliar, possibly reflecting awareness of musical expectancy (musicians' recognition better than non-musicians). Although there was an expectancy for correct phrase endings in both domains, identified differences (language semantic expectation processes [N400] and musical semantic expectation processes [P600]) possibly arose due to an absence of understanding of the music's fundamental meaning. Although both melodies were instrumental, the familiar melody did contain associated lyrics (Toreador's song from Carmen by Bizet) that may have affected the results. However, some processing specificity for semantic aspects of music and language was shown.

When semantic violations are combined with music syntactic incongruities non-neurological tasks have identified interference effects. For example, in a lexical decision task (Poulin-Charronnat, Bigand, Madurell, & Peereman, 2005), semantically related and unrelated sentence final words were sung to either a tonic chord (expected) or a less referential chord (subdominant). Results showed a significant interaction between semantic and harmonic relatedness. This suggests that, in vocal music,

---

[5] Normal response to words, negative at 400ms from word onset.
[6] Response to harmonic syntactic violations, positive at 400-1,000ms .
[7] P600 amplitude (response to syntactic violations, positive at 600ms from word onset).

harmony has the potential to interfere with language processing. In addition, Slevc, Rosenberg, and Patel (2009) showed slower reading times when accompanied by out-of-key chords and semantic ambiguity suggesting some interference effects from the combination of semantic violation and syntactic anomalies. Sentence reading with syntactic anomalies, with and without concurrent listening to related pitch patterns "clusters", suggests the shared syntactic processing of language and music are related more to reintegration rather than to the rules of syntactic combination (Van de Cavey, Severens, & Hartsuiker, 2017), thus supporting overlapping processing (Patel, 2003). In other words, when highly distracting semantic violations (unrelated words) and music syntactical incongruities (sung to an unexpected chord) are combined (Kunert et al., 2015; Perruchet & Poulin-Charronnat, 2013; Poulin-Charronnat et al., 2005; Van de Cavey et al., 2017) the semantically-based verbal effect and semantically-based music effect have the potential to give rise to increased disruption. (See Chapter II for violation study discussion.)

1.3 Musical semantic memory

As a means of communication music and language convey meaning because they are able to create an expectation for what is to follow (Meyer, 1956). When semantic expectation is realised by the following note, chord, word, there is contentment, when unresolved, or not according to "rule", there is surprise or even tension. However, considering the disparate cognitive operations necessary to process these systems of communication, musical grammar has been shown to react in a similar way to language grammar in relation to rule-based knowledge (Besson et al., 1998; Miranda & Ullman, 2007). However, this similarity is not evident when experiencing semantic incongruities.

Familiarity is considered to be the main index of musical semantic memory and is judged according to whether a melody "felt" familiar (Platel, Baron, Desgranges, Bernard, & Eustache, 2003, p. 244). Two delayed tasks presented to identify recognition of familiar and non-familiar melodies in order to judge musical episodic memory (Platel et al., 2003) have shown increased cerebral blood flow in different brain areas for episodic (melody recognition irrespective of familiarity), and semantic musical memory (well-known, stored in a proposed musical lexicon). Hence there is empirical support for the notion of two different memory systems (episodic and semantic) in music memory.

1.3.1 Familiarity *versus* unfamiliarity.

Recognition of a stimulus as familiar or unfamiliar, requires diagramming various components associated with the stimulus that combine into a unified recognizable composition. Relevant areas of LTM are then activated. Specific interaction and combination between pitch contour and rhythm have been identified as the most effective cues for identification/recognition of a familiar melody (Hébert & Peretz, 1997). Recall of verbal or musical items requires recognition, but not necessarily of their entirety. Passages of language comprehension involve retaining "gist" rather than exact order of words within sentences. Likewise, for recognition of a known melody arguably, only the "gist" will be identified (Dowling & Harwood, 1986) rather than discrete musical detail (such as tempo, dynamic, and instrumentation). This would account for melody recognition, for example, in a different key. However, this relatively abstract memory also requires surface characteristics to be employed that contribute to the melody's uniqueness (Raffman, 1993).

For music to be recognised as familiar it needs the auditory stimulus to contact, and then be selected, from all items of musical patterns that exist in memory. This musical processing memory store has been termed the "musical lexicon—a perceptual representational system for isolated tunes" (Peretz et al., 2009, p. 257; Peretz & Coltheart, 2003; Peretz & Zatorre, 2005) containing representations of one's complete musical experience. It is thought to be accessed automatically and consists of three stages (Peretz et al., 2009). First, an "access" stage begins when music is detected, second, "selection" of relevant item from all in the lexicon (this stage can also produce singing from memory, or inner singing). Third, "integration", whereby each item is set in context. Here, rules of syntax and additional musical associations could be integrated. The model postulates existence of a verbal and musical lexicon, independent but having a privileged relationship with one another (Gröussard et al., 2010). Yet, if this system is unique to music processing, or if it can be co-opted into processing other sounds, creating an overlap with sounds such as speech, is less clear (Gröussard et al., 2010).

Using familiar and non-familiar melodies, Platel (2005) hypothesized, following previous neuropsychological literature (Halpern & Zatorre, 1999) and PET studies (Platel et al., 2003; Platel et al., 1997), that right temporal areas, in addition to left temporal areas, would be activated by musical semantic processes. His task contained four conditions: one semantic, one episodic, one perceptual control, and a rest measurement. Music material consisted of 64 familiar and 64 unfamiliar 5-second actual flute melodies (with no word association). In the semantic condition melodies were identified as familiar or unfamiliar. A following task (episodic condition) required recognition of previously heard melodies (with distracters). Perceptual control focused on identification of the last two pitch notes rated similar or different. Another set of familiar and unfamiliar melodies were used in a control test. Two separate patterns of activation were identified. Results showed left hemisphere dominance for semantic memory with right hemisphere dominance for episodic, supporting previous literature (Platel et al., 2003; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994). However, Platel (2005) did suggest some functional specificity for music semantic memory based on partial overlap in temporal activation for melodies with those obtained for phonemes or words.

More recently, Transcranial Direct Current Stimulation (tDCS [8]; Schaal, Javadi, Halpern, Pollok, & Banissy, 2015) has been used to detect melody familiarity. Anodal current stimulation to the right posterior parietal cortices showed a marked deterioration for melody recognition (new from previously memorized, following a 7-minute retention period), as opposed to stimulation to the left side, or without direct current stimulation. Results indicated a right lateralization for melody (Schaal et al., 2015). However, areas responsible for specific stages, encoding or recognition, of melodic memory processes were not clarified.

Successful recognition of a familiar melody, or acknowledgement of an unfamiliar melody, may not necessarily utilize the same processing components. To compare cerebral responses to random tones, familiar melodies (no lyrics association), and unfamiliar melodies (retrograde familiar melody—played backwards, see Hébert & Peretz, 1997), participants were scanned using fMRI (Peretz et al., 2009). Judgement was made according to a 1-10 level of familiarity. Familiar melodies were found to activate the right superior temporal sulcus a “critical region involved in music processing” (p. 256) supporting earlier work by Plailly, Tillman, and Royet (2007). Furthermore, auditory memories were closely linked to singing, by activation in areas such as the supplementary motor area that possibly relate to subvocalization (melody singing has been found to be a more reliable indicator of familiarity than title retrieval [Dalla Bella, Peretz, & Aronoff, 2003]). Although the task melodies of Plailly et al. (2007) had no lyrics association, their tuneful contour appeared to promote covert singing. The forthcoming empirical work in the current thesis will require overt production of a familiar melody/lyrics from LTM during presentation of a different

[8] tDCS sends low level current through electrodes attached to the scalp which induces intracerebral current flow and allow brain waves to be traced.

irrelevant melody/lyrics with associated and novel lyrics/melody. According to Plailly et al. (2007) the supplementary motor-area should be differentially activated according to associated familiarity. This issue is returned to in Chapter III-IV.

1.3.2 Timbre and tempo.

Language and music processing have been compared mainly at semantic and syntactic levels. However, timbre and tempo change can also impair explicit, conscious, memory for target material (Halpern & Müllensiefen, 2008). Differences in memorization for verbal and musical items was noted by Halpern and Müllensiefen (2008) when item timbre or tempo was changed. Implicit musical memory (knowing how to e.g., play an instrument) was measured by rating "pleasantness of old and new melodies" and explicit memory (conscious, contextual e.g., ability to recognise a previously heard melody) measured by confidence recognitions rating. Half of previously heard excerpts were modified by timbre or tempo change. Findings revealed that timbre and tempo change impaired explicit memory but only tempo changes impacted detrimentally on implicit melody recognition. Halpern and Müllensiefen (2008) therefore concluded that both implicit and explicit memory are used to process music.

To support the brain in identifying relevancy of language and sound, their vocal tone colour (timbre) classifies those specific characteristics which enable sounds/words to be identified as relevant or irrelevant, and so affects how a sound is processed (Trollinger, 2010). Perceived unimportant sounds seem to be processed in right temporal areas as opposed to perceived meaningful sounds processed in left areas. Furthermore, it appears, from findings following studies of people with left temporal lesions, that implicit retrieval of melodies can still be preserved despite disrupted explicit retrieval abilities (Peretz, Gaudreau, & Bonnel, 1998; Samson & Peretz, 2005). However, implicit (melody retrieval, as measured by liking judgements) and explicit (melody recognition) tasks were impaired in people with right temporal lesions. Samson and Peretz (2005) concluded the right temporal lobe's function is to support melody recognition through creation of melody representations, whereas the left-sided temporal lobe is more involved in explicit retrieval of melody. However, as there appears to be similar brain behaviour when reading aloud or singing (Trollinger, 2010), any overlap between music and language processes should be most apparent through the medium of song.

1.4 Vocal music

1.4.1 Integration.

A coming together of melody and speech witnesses two unique human skills in action. Vocal music is developed from two components, words (lyrics) and tune (melody). Yet studies exploring whether these two components are integrated or independent of one another have produced conflicting results. Previous studies into the relationships between lyrics and melody in song show melodies of songs are better recognised when heard with their original words than when heard with text of another equally familiar song (Bonnel et al., 2001). Similarly, words of songs are better recognised when heard with their original melody than when heard with a different but equally familiar melody (Crowder, Serafine, & Repp, 1990; Samson & Zatorre, 1991). However, when listening to song it is unclear if these two components are processed together or separately (Hamzelou, 2010; Hébert & Peretz, 2001).

When learning song lyrics, Broca and Wernicke's areas of the left hemisphere are engaged, for learning melody the right hemisphere is more active, but "singing activates the whole brain" (Trollinger, 2010, p. 21). Schlaug, Marchina, Norton, and Wan (2010) showed singing engages right-hemisphere fronto-temporal regions, allowing left and right hemisphere connections to be formed. For example, a stroke patient with damage to the left side of their brain, which affected their ability to form words, was taught to communicate through song (e.g., to say "I am thirsty"). In this context, the music appeared to stimulate the right hemisphere thereby facilitating crucial connections (Schlaug, Marchina, & Norton, 2008, p. 316). Schlaug et al. (2010) concluded, therefore, that singing might be an alternative medium to re-engage brain areas.

Although processing pathways for melody and associated lyrics are still unclear there is an accumulation of data from recognition studies (Serafine, Crowder, & Repp, 1984; Serafine, Davidson, Crowder, & Repp, 1986) and neuropsychological studies (Hébert & Peretz, 1997; Samson & Zattore, 1991) suggesting integration. For example, Serafine et al. (1984) hypothesised integration (melody with lyrics) following recognition of folk song melodies and texts both old (e.g., previously learned and therefore within LTM) and new (presumably held with STM or episodic memory). Their experimental manipulations consisted of pairing a novel melody to existing text or a novel text to an existing melody across five conditions: i old song, ii new song, iii old melody new lyrics, iv new melody old lyrics, v old melody old lyrics from a different song. Recognition of original pairings of melody and lyrics was highest, as opposed to a different pairing, suggesting integration across different performance conditions. As recognition of one element was facilitated by simultaneous presence of the other this argues for integration. If independent, equivalence of melody/text recognition would have been evidenced, and if holistic storing, one component would have been recognised only with its original partner. Their findings demonstrated better recognition for exact songs (no manipulation) as opposed to mismatched songs (manipulation of either melody or text). This could indicate that pairing to-be-remembered information with music facilitates its recall, whereby music may act as a retrieval cue to evoke to-be-remembered information (Rainey & Larsen, 2002).

Few neurological studies have addressed cognition for pairing of melody and lyrics within song (Besson et al., 1998; Bonnel et al., 2001; Peretz et al., 2004) and differing results for integration or separation of these two components have been obtained from fMRI and ERP studies from song production. For example, Schön et al. (2010) focused on song perception and identified extensive music and language interactions following listening to pairs of spoken, vocalized (sung on the syllable “vi”), or sung words in a same-different task. Processing of both words and music was necessary in the song condition as either could be different. Both hemispheres were activated by sung or spoken words, although not to the same extent (right temporal areas showed increased activity for music processing). Therefore, Schön et al. (2010) argued against specificity for music and language suggesting lexical, phonological, and music processing are located in a common cerebral network.

1.4.2 Independence.

Disagreement regarding the superior/equal/partnership between melody and lyric within vocal music is longstanding (Boulez, 1966). Melody, especially when combined with lyrics in song, comprises multiple components that affect multiple brain regions in both right and left-brain hemispheres (Gordon, 1981; Peretz, 1993). Whilst some prior research promotes the view that music and lyrics within song are integrated in memory (Serafine et al., 1984) other neuroimaging research suggests independence (Besson et al., 1998), or that distinct, shared, and overlapping cognitive resources exist between music and language (Koelsch, 2013; Patel, 2008).

To ascertain how melody and lyrics are processed when listening to song, a comparison between the semantic aspects of language (semantically congruous or incongruous final words of operatic songs) and the harmonic (syntactic) aspects of music (sung in or out-of-key) has been addressed in ERP studies (e.g., Besson et al., 1998). If melody and text are integrated (Serafine et al., 1984) an interaction would be shown between N400[9] produced by semantic violation and P300[10] produced by music violation. If processing-independent, ERPs should be additive (double violation equal to the sum of both individual violations, see Sternberg, 1969). Besson et al. (1998) recorded responses to French a cappella operatic song excerpts with melodic and semantic violations to final words (ending with congruous/incongruous words sung in/out of key) manipulated across four conditions. Besson et al.'s results confirmed N400 components for semantic incongruities (which has been identified for written or spoken speech [Kutas & Hillyard, 1980]) to be generated in the same way for sung or spoken passages thus showing musical structure had no effect on semantic processing. Similarly P300 associated with syntactic incongruities, were identified with out-of-key sung-words and out-of-diatonic scale chords, with no statistical difference shown in the additivity tests. Besson et al. (1998) concluded independence of melody and lyrics since pitch/harmonic processing was unaffected by lyric semantics. These results were replicated in an identical condition task but with analysis of correct responses (Besson et al., 1998). Together such findings support processing independence.

Following their findings supporting separation for semantic aspects of music and language, Besson and Schön (2001, Experiment 2) sought to determine if the N400 found with spoken words (Kutas & Hillyard, 1980) could be replicated with sung-words, or if the P600, found with melody, would be identified from a group of musicians. Four conditions (semantically congruous, sung in tune; semantically incongruous, sung in tune; semantically congruous, sung out of tune; semantically incongruous, sung out of tune) derived from a cappella French opera excerpts. With a participant group of musicians they found sung incongruous words resulted an N400 (as with language), and incongruous (out of tune) sung words elicited a P600 (as with melody). Moreover, a double incongruity test, words sung out of tune showed N400 and P600 components. As N400 arose before P600 they suggest that although lyrics and melody were processed independently the language component was processed first. Their previous results were replicated. An N400 identified for sung incongruous words with a P600 for congruous words sung out of tune. Interestingly, N400 was eliminated when participants focused on the music and P600 eliminated when concentrating on words. Musicians inability to disregard melodic incongruities evidencing independent processing of lyrics and melody.

---

[9] N400 (normal response to words, negative at 400ms from word onset).
[10] P600 and LPC produce an outcome known as P300 (Besson et al., 1998).

The independence view has been further supported by findings from additional non-neurological tasks that also aimed to detect melodic and semantic incongruities in French operatic songs (Bonnel et al., 2001) using a dual task methodology whereby the single task was to focus on language or music, and the dual task demanded divided attention to identify inappropriate language and music. In this setting, recognition of final words, from complete phrases of French operatic songs, as semantically incongruous and whether sung out of the home key was required. Incongruous final words rhymed with the original, incongruous pitch being a subtonic (non-diatonic) below the tonic thereby violating expectation. Performance levels were similar across both tasks. Irrespective of participant's musical expertise, based on a 6-point scale self-reported judgement of correctness, participants were equally able to detect a violation that was music or language based in the single and in the dual task (although musicians were more accurate). In dual tasks the nature of the task is influential, with identification of the violation being critical. Here, however, only detection, not identification, of the violation was necessary.

1.4.3 Melody and text association.

Song is routinely learnt through performance and musical interaction, therefore, cues become deeply embedded (Rainey & Larson, 2002). Moreover, by repetitive learning of a familiar melody/lyrics a link is developed between mental representation and motor production (Dalla Bella, Berkowska, & Sowiński, 2015; Tsang et al., 2011). Speculatively, if familiar words/melody are strongly connected in LTM a familiar melody/lyrics could help retrieval of associated lyrics/melody. Finding novel spoken texts can be encoded along with melody, then, better recognized following re-presentation of that melody, suggests text learning can occur through association between text and melody (cf. Crowder et al., 1990). Moreover, Crowder et al. (1990) found melody learning leads to mandatory retrieval of associated text that appears to be based somewhat within the perceptual rather than semantic system. For example, Crowder et al. (1990) reported a melody better recognized when repeated with nonsense words phonologically similar to nonsense words with which the melody was originally presented. Although findings are compelling, these recognition tasks did not require production, therefore, cannot tap the extent to which retrieval processes involving long-term (e.g., semantic) memory or speech planning may be facilitated by melody. Such will be an important aspect for the upcoming experimental series. Processing specificity between melody and language (lyrics), specifically with regard to musical semantic memory (familiarity), and if musical training is influential on task outcome, will be tested in this forthcoming series of studies through a different medium—distraction—(see Chapter II).

1.5 Listening and production

1.5.1 Musicians *versus* non-musicians: Training.

Experience, musical training, and the chosen medium of output are factors for consideration in music related research experiments, especially when experiments, such as Study I in this series, require musical performance. Arguably, musically trained participants have advantage in music tasks not necessarily replicated for verbal tasks (Cuddy & Cohen, 1976; Williamson, Baddeley, & Hitch, 2010; but see Jakobson, Lewycky, Kilgour, & Stoesz, 2008; Zuk, Benjamin, Kenyon, & Gaab, 2014). However, definition of a musician lacks specificity. An individual's equivalence in proficiency between language and music would seem to occur following seven years of instrumental/vocal training, four being consecutive (Ericsson, Krampe, & Tesch-Römer, 1993). This standard is generally accepted and some studies have differentiated between musically and non-musically trained participants (see Chapter II). In the studies reported within the forthcoming empirical series of this thesis, participants were not recruited according to their musical ability, but consideration of their musical abilities in relation to task performance in the presence and absence of distracters was investigated.

1.5.2 Musicians *versus* non-musicians: Memory.

There is some evidence from neuroimaging studies, that musicians, as compared to non-musicians, use two WM systems (Schulze et al, 2011). Neural differences associated with musical training identified from brain-imaging studies have reported enhanced cognitive abilities in musicians, such as verbal memory (Ho, Cheung, & Chan, 2003) and WM (Berti et al., 2006). Although core structures, such as Broca's area and SMA, are utilized by both groups, for musicians, it appears that specific neural subcomponents of WM are utilized only for verbal information (right insular cortex), or only for tonal information (right globus pallidus, right caudate nucleus and left cerebellum; Schulze & Koelsch, 2012; Schulze, Mueller, & Koelsch, 2011). The proposed two WM systems purportedly comprise the phonological loop (e.g., Baddlely, 1986)—for rehearsing verbal information—and a tonal loop—supporting pitch rehearsal. Strong links are forged between production and auditory WM, based on a musician's experience of how to produce sound, involving sensorimotor representations to maintain auditory information of target items in WM (Schulze & Koelsch, 2012;

Schultz et al., 2011). In non-musicians an overlap between the professed phonological loop and possible existence of a tonal loop, as a subsystem of the articulatory loop, has been identified via fMRI studies (Hickok, Buchsbaum, Humphries, & Muftuler, 2003).

In unpublished work (Jordan, 2017) explored retention of tone sequences under singing suppression—la-la-la-la-la on an ascending/descending major triad—that required use of the proposed tonal loop for tone sequence pattern rehearsal) and articulatory suppression (repeated “the” reliant on the verbally-based phonological loop). If a special tonal loop for musicians exists, singing suppression should cause greater disruption for musicians, while non-musicians should be more greatly impaired by articulatory suppression. Jordan (2017) found only singing suppression impaired musicians’ recall of tonal sequences, thereby supporting the idea of a tonal-loop. However, although articulatory suppression did impair non-musicians’ recall, singing suppression caused the greater disruption thereby suggesting non-musicians also have some access to the purported tonal loop within WM. Furthermore, in a subsequent atonal same/difference detection task, although singing suppression was more disruptive to musicians than articulatory suppression, non-musicians were greatly disrupted by both suppression activities. However, dodecaphonic patterns are unexpected and difficult to comprehend, empowering the musician with an advantage based on musical experience, tactical understanding, and greater reliance on the proposed tonal-loop.

Musicians’ ability to process some auditory features, such as tonal pitch, develops according to processing familiarity (Marques, Moreno, Castro, & Besson, 2007). Practice makes perfect. This honed skill was identified in a task to compare sequences of two tones as same/different, separated by the presentation of different tones, verbal, and visual material during inter-stimulus interval (Pechmann & Mohr, 1992). Although there was equivalence between musicians and non-musicians when presented by tones, only non-musicians were also affected by the irrelevant verbal/visual items (Pechmann & Mohr, 1992). This suggests that although dissimilar distraction (verbal/visual items) created an interference effect for non-musicians, musician’s familiarity with tonal nuance modulated the effect. In addition to linguistic and music task superiority (Fitzroy & Sanders, 2013; Jentschke & Koelsch, 2009; Schön et al., 2004), there has been recent interest in whether musician’s musical memory can be replicated for other types of memory. Changes in neuronal circuity that ensue from high levels of musical motor skill development, usually from an early age, may impact other areas of the nervous system (Watson, 2006).

Currently, growing literature shows better performance by musicians in tasks requiring higher cognitive demands, improved auditory attention, and executive functioning (Zuk et al, 2014). For example, Jakobson et al. (2008) compared the results of a 16-word recall task, from four semantic categories, by musically and non-musically trained students. Musician scores were higher. A second task followed the same procedure but word recall was replaced by recall of two geometric shape pictures (variations of a square and a line). Again, musicians outperformed non-musicians. Speculatively, musical training develops enhanced encoding strategies during learning, utilized in performance, such as an ability to chunk items into related groups when analysing a musical score (Halpern & Bower, 1982).

It would appear that musician's experience allows them to adopt an analytical method of processing music with enhanced ability to use contour or interval information flexibly (Peretz & Babai, 1992). For example, musicians may employ different strategies encoding pitch, such as tonic sol-fa system (doh, maw, lah), intervallic description (minor 3rd, augmented 4th), or letter name (C, E*b,* A). Taken together, these findings suggest that the multidimensional tasks of word-learning and music performance, requiring auditory perception, prosodic processing, focused attention, and memory, could be processed in interaction with other cognitive functions.

1.6 Production of musical sound: Vocal

Probably the oldest and universal method of learning and performing music, with or without words, is by singing, in existence from infancy and requiring no musical training (Dalla Bella et al., 2015; Peretz et al., 2004). Singing has been shown to activate the "whole brain area" (Trollinger, 2010, p. 21) and engage perceptual, motor, and sensorimotor processes (Dalla Bella et al., 2015). Pathways of information flow for familiar song output from long-term memory, as compared to a STM repetition task, may be considered with reference to the modular model of music processing that details a route through the musical and verbal lexicons to vocal output (Peretz & Coltheart, 2003, see Chapter IV, 4.1.1). Although, it is unclear how these pathways adjust when the output requirements change, such as when required to speak or sing/hum as required for this series of study.

According to the vocal sensorimotor-loop model (Dalla Bella et al., 2011), a constant comparison between output of vocal production and internal representation of to-be-performed sound is needed (Tsang et al., 2011). Production of a vocal sound requires memory of the sound to be reproduced. Such reproduction entails motor skills necessary to produce the sound, perception, and feedback relating to the outcome of the emitted sound (Tsang et al., 2011). Therefore, when required to sing, various disparate process stages are undertaken. With regard to the vocal sensorimotor loop singing model (Dalla Bella et al., 2011) this process begins with retrieval from memory of various musical elements to be incorporated in the performance, such as pitch and duration. These are directed to motor control to facilitate sound output in the form of singing. Finally, there is a process of auditory feedback to allow for output comparison and correction. A combination of vocal and motor skills required for accurate pitch production, developed over time, are crucial for the accuracy of melody output. In the same way that instrumentalists develops finger memory, singers form habits for lungs, voice box, mouth, lips, and tongue. Wise and Sloboda (2008) identified individuals considered to be unable to sing in tune (acknowledged as tone-deaf) but nevertheless well able to respond to short melodic phrases, especially if repeated or accompanied. Further, Dalla Bella, Giguere, and Peretz (2007) report that non-trained singers have accurate memory for pitch and tempo. These findings raise the question as to whether difference in production/performance of vocal music results from ineffective initial encoding of auditory information and ability to compare this with motor output (Gillis, 2014). According to Hutchins and Peretz (2012) there are multiple causes of poor singing not just poor perceptual ability. They concluded, following greater accuracy from non-musicians when matching pitch using a slider than with their voice, that the greatest inhibitor was poor motor-control. A prerequisite for accurate musical performance, irrespective of production vehicle, singing, humming, or instrumental, is therefore the ability to hold and "hear" a musical phrase in memory for comparison with future motor output.

1.7 Auditory imagery

Melody, words, or indeed environmental sounds do not need to be audible to be "heard". The notion of "seeing with the mind's eye" (Kosslyn, Ganis, & Thompson, 2001, p. 635) has been influential to our understanding of mental imagery. It has become acknowledged a partnership exists between "inner ear" and "inner voice" necessary for retaining verbal items in STM (e.g., Gathercole & Baddeley, 1993). The ensuing interplay is then enhanced by analytical judgements made concerning the auditory stimuli (Baddeley, 1986, 1990; Gathercole & Baddeley, 1993; see section 1.2.1). Speech items stored in LTM can become reactivated by a perceptual experience that does not require overt stimuli, in other words, produced during mental imagery

(Kosslyn, 2005; Tian & Poeppel, 2012) whereby a link is formed with the motor systems. When listening to speech, the continuous acoustic stream is parsed into separate items (e.g., syllables or words) to allow comprehension. Comparable with the duck-rabbit changing visual image (Warren & Gregory, 1958) rapid repetition of words can result in an ambiguous sound stream causing the verbal transformation effect (e.g., life…fly, Warren, 1961). Covert speech has also been shown influential in tasks requiring auditory imagery (Smith et al., 1995).

Auditory imagery, utilizing the "inner ear" (Aleman, Nieuwenstein, Bocker, & De Hann, 2000) enables sound to be heard without overt aural activity. In addition to the speech-planning mechanism being opportunistically co-opted for retaining visually presented to-be remembered items (Jones et al., 2004), it is also a medium through which melody can be internally generated. Moreover, how we listen to and imagine music can produce the same neural activity (Halpern & Zatorre, 1999; King, 2006). For example, participants reported the experience of hearing familiar, but not unfamiliar, popular songs continue during various interpolated periods of silence (Kraemer, Macrae, Green, & Kelley, 2005).

Based on the WM model (Baddeley, 1986; Baddeley & Wilson, 1985; Gathercole & Baddeley, 1993) interplay between inner ear (store) and inner voice (sub-vocal rehearsal) has been examined via subvocalization. For example, participants were requested to use a continuous irrelevant utterance (tah) that impaired judgements as to the rise or fall from 2nd to 3rd note within a familiar melody when instructed to imagine it from a song-title (Smith et al., 1995). Results showed the inner-ear/inner-voice partnership (subvocalization) was relevant when required to present/analyze auditory items, such as a familiar melody. Smith et al. (1995) concluded that the phonological loop was necessary when cognitive judgements are required (such as "parsing meaningful letter strings" and "scanning familiar melodies"). Whereas mental judgements about "pronunciation of word-final phonemes" (e.g., words ending in 's' pronounced 's' or 'z') whilst overtly repeating "Suzie-Suzie", and "homophone judgements", whilst repeatedly counting numbers one to six as articulatory suppression, required no inner-ear/inner-voice partnership arguably due to a direct access route from visual/grapheme to phonology.

That articulatory suppression and task-irrelevant sound disrupts judgements of rising and falling pitch within familiar melodies suggests that articulatory planning is required to generate the auditory imagery required for the judgement task (Smith et al., 1995). However, this says nothing about how lyrical retrieval might be impacted upon by articulatory suppression or task-irrelevant sound. Moreover, Smith et al. (1995) used only familiar melody as a distracter and thus did not compare melody with/without lyrics, or that was familiar/unfamiliar on target melody or lyrics retrieval leaving the potential role of these variables unknown. Further, motor system requirements for recognition may not be as demanding as for production tasks. These omissions will be addressed later in the content of the empirical studies reported.

1.7.1 Role of motor-systems in auditory imagery: Perception and action.

Perception of music involves the human motor system in action simulation (Leman, 2007). The embodied music cognition theory suggests music perception activates motor systems that allow melodies to be reproduced or performed (see Chapter III, 3.1.4; Leman, 2007; Leman & Maes, 2015). Therefore, with regard to this series of studies, the production of a familiar target song from LTM and the mere perception of a different familiar to-be-ignored song should involve some degree of activation within the motor system. Moreover, since familiar melodies, as compared to unfamiliar melodies, are well-represented in LTM such motor activation should be greater. Subvocalization (covert singing of lyrics using inner speech) is also implicated in rehearsal of auditory images (Reisberg, Smith, Baxter, & Sonenshine, 1989). Another musical feature, intensity (dynamic) imagery has also been explored (Bailes, 2002; Bailes, Bishop, Stevens, & Dean, 2012). Following findings from a study with musicians Bailes et al. (2012) suggested a perceived relationship between the dynamic of the music "heard", and the force required for production, inferred a motor representation. Therefore, although a person's ability to produce sound via subsequent motor action and to perceive/imagine sound varies considerably, the overriding issue for this impending series of studies is that auditory imagery may be influential in recall of melody/lyrics from previously-encoded song.

1.7.2 Role of motor-systems in auditory imagery: Inner ear.

Unintentional auditory/musical imagery may be readily identified by "earworms" that relentlessly invade our thoughts (Beaman & Williams, 2010; Liikkanen, 2009). Described as an involuntary memory triggered by specific identifiable cues (Morton, 1990), once in control our ability to dislodge the so called "sticky music" (Sacks, 2007) and prevent the ensuing involuntary cognition is limited. Beauman et al. (2007) identified similar brain regions are activated when imagining or performing music. According to McCullough Campbell and Margulis (2015) the motor cortex reaction to music contributes to the earworm effect, with motor planning brain regions being fully active even when music is merely imagined. Therefore, the ability to exercise necessary motor-control to produce components of familiar song during involuntary motor-activity produced by irrelevant familiar song, overt or covert, will be tested in the forthcoming empirical studies. From the foregoing description, it follows that the nature of the irrelevant stimuli may be important in driving an earworm effect. Familiar nursery rhymes, due to their repetitive musical features, should activate the motor-system and thus their impact of this competition for motor-control will be investigated in the context of retrieval of target melody/lyrics (Study I-II) and rehearsal (Study III).

Although reaction to an earworm is very individual, they do all contain certain musical characteristics (Jakubowski, Finkl, Stewart, & Müllensiefen, 2016). Jakubowski et al. (2016) identified a fast tempo, a repetitive motive or riff (e.g. the opening of "Moves like Jagger"), and an easy to remember melodic line frequently with a rising first phrase contour answered by a falling second phrase (e.g. "Twinkle, twinkle, little star", inherent characteristics exploited in to-be-reported studies). Jakubowski et al. (2016) observed earworms contain a unique feature (e.g. 5/4 time of Dave Brubeck's "Take Five"), an unexpected leap (e.g. tritone in Bernstein's "Maria" from West Side Story, or "In the Mood" by Glen Miller). However, involvement in conscious distraction, such as by another music or language task, or by engaging with the sound (Williamson et al., 2011), or chewing gum (Beaman, Powell, & Rapley, 2015), seems to reduce an earworm's impact. Following the suggestion that jaw movement interferes with STM, Beaman et al. (2015) identified how gum-chewing (compared to non-chewing) reduced earworm recollection of catchy songs, participants had been instructed to ignore. Chewing also impacted on ability to "hear" the imagined music. Chewing impairs use of the articulators and the diminishment of earworms while chewing therefore suggests that earworms draw upon the same articulatory and phonological system upon which STM is reliant.

Item recall in exact order of presentation—seriation—has been acknowledged a prerequisite for task disruption via suppression (see Chapter II). Previous studies have reported subvocalization blocked by chewing candy, resulting in STM item order and item recognition deterioration (Kozlov, Hughes, & Jones, 2012). It should be noted however, that it would appear surprising that the missing-item task (see Chapter VI) that, arguably, does not require serial order recall and known to be unimpaired by articulatory suppression (Beaman & Jones, 1997), can be impaired by chewing gum. One possible explanation for this outcome (Koslov et al., 2012) is that disruption of the missing-item task could result from a proportion of participants that adopt item rehearsal as the mnemonic strategy for task recall (Morrison, Rosebaum, Fair, & Chein, 2016).

1.8 Music as a mnemonic

A musical mnemonic can support learning of non-musical items by pairing those items with musical elements, such as pitch or rhythm, to support rehearsal in memory (Gardiner & Thaut, 2014; Wallace, 1994). The idea that memory for musical tones and verbal items are processed by similar WM processes (Pechmann & Mohr, 1992; Salamé & Baddeley, 1989) has been reinforced by recall facilitation of information when music is used as a mnemonic (Jellison, 1976; Moussard, Bigand, Belleville, & Peretz, 2012; Sloboda & Parker, 1985; Wolfe & Hom, 1993), and when retention of words has been supported by music, such as when learning a foreign language (e.g., Bhatara, Yeung, & Nazzi, 2000). Although lyrics appear to better cue melody recall than vice versa, overall melodies are better recalled than lyrics (Peynircioğlu, Wagner, Baxter, & Shaffer, 2008). With a familiar melody, novel words learnt as song were "more strongly integrated in the mental lexicon compared to words learnt in the spoken modality" (e.g., Tamminen, Rastle, Darby, Lucas, & Williamson, 2015, p. 3). Furthermore, improved unfamiliar text recall when studied as song, rather than spoken, shows a musical association during learning facilitates text retention (Palisson et al., 2015; Rainey & Larson, 2002; Simmons-Stern, Budson, & Ally, 2010).

Empirical evidence that a music mnemonic may support communication for people with cognitive impairment, is limited but has afforded some investigation. Through aging, ability to retain declarative information can become impaired although procedural memory can be largely unaffected. Worldwide 46.8 million people were living with dementia in 2015. This is projected to rise to 74.7 million in 2030, and to 131.5 million by 2050 (Alzheimer's Disease International, 2015). However some people with semantic dementia have preserved knowledge of music (Hailstone, Omar, & Warren, 2009; Weinstein et al., 2011). Therefore, developing potential benefits of memory for song (especially those learnt in early life, Krumhansl & Zupnik, 2013), and its use as a mnemonic, may support increased communication and interaction while pharmaceutical research continues (Hara, 2011; Johnson et al., 2011). AD is characterised by progressive decline in cognitive functions that present, in the early stages, as impaired episodic memory of events and experiences, but seems to leave procedural memory—how to perform actions—largely intact (e.g., Baird & Samson, 2009). Music association can support memory retention (Ally, 2010). For example, Palisson et al. (2015) measured the learning of texts with either sung, spoken alone, or spoken to a silent film clip, immediately and after a 5-minute delay, from people with AD. Findings suggest a musical association during learning facilitated text learning and retention. Simmons-Stern et al. (2012) reported improved content knowledge in controls and AD participants when matching melodies (children's songs) with objects and related actions (e.g. pills/fill the pillbox with your pills). Although they were unable to support enhanced memory for specific content information, preserved musical memories through aging have been reported from case studies. For example, recognition of familiar, as compared to unfamiliar/novel, melodies, and recognition of incorrect pitch (based on "sing along" response and facial grimaces) were identified from an 84-year-old woman with diagnosed dementia (MMSE score of 8[11], Cuddy & Duffin, 2005).

To capitalise on this purported music memory retention through aging/brain damage, and to develop music's potential as a mnemonic, techniques that support increased sustainable interaction, such as singing, are necessary (Heathcote, 2009). Of particular interest is that music can synchronize brain responses across listeners and performers (Abrams et al., 2013; Sänger, Muller, & Lindenberger, 2012). Using scalp attached electrodes, the brain waves of two guitarists synchronised when performing the same and different parts (Sänger et al., 2012). The emergence of interbrain networks during music listening/performing could, therefore, potentially support communication through age related cognitive decline (Astell & Ellis, 2006; Götel, Brown, & Ekman, 2000, 2008). Speculatively, if strongly connected in memory, producing familiar melodies may help retrieval of associated lyrics, although these issues are far from resolved (Peretz, Radeau, & Arguin, 2004b).

Any suggestion that memory for music may be preserved whilst other language-based cognitive faculties such as verbal fluency becomes impaired (Cuddy et al., 2012; Simmons-Stern, 2010, Simmons-Stern et al., 2012), again highlights the integration (Serafine et al., 1986), independence (Peretz & Zatorre, 2005) question of music and language within song. If integration is established, there could be potential benefits for population groups who have been found to recall words to song better than spoken words (Moussard et al., 2012; Prickett & Moore, 1991), show improved memory for information when paired with a familiar melody (Wolfe & Hom, 1993), or with an unfamiliar melody constantly repeated (Calvert & Billingsley, 1998). Arguably, this

[11] MMSE scores 30-25 suggests normal cognitive functioning, 20 to 24 points suggests mild dementia, 13 to 20 suggests moderate dementia, and less than 12 indicates severe dementia (Alzheimer's Society).

may result from chunking (Jones, 2012). This potent aid to remembering is described as “the very lifeblood of the thought process” (Miller, 1956, p. 95).

1.9 Interim Summary

Memory for song may be regarded as special in that all musical elements pitch, duration, dynamic, tempo, texture, timbre, and emotion, coupled with a language component, lyrics, are encoded together. Much of the current debate centres around verbal and musical integration or independence, and if music can help facilitate acquisition of new information. During the 1970’s and 80’s cognitive issues, such as memory, attention, and musical syntax, were central to research into the mental world of focused music listening (North & Hargreaves, 1997). More recent research, however, is cognisant of an interaction between cognitive factors and real-life behaviour involving passive music listening (North & Hargreaves, 2008). Many distraction studies consider these issues by addressing recognition of pitch/phrase/word in STM tasks, but recall through performance, using complete familiar music material, may challenge the dominant view pertaining to specialization for neural mechanisms of music processing. Therefore, to address this gap, the empirical studies undertaken in this thesis will adopt novel tasks that aim to identify similarities and differences in retrieval patterns of the major components of song, melody and lyrics, produced via performance during the presence of potentially distracting musical and non-musical task-irrelevant sound. By exploring the role of concurrent auditory distraction on performance (Study I-II), vulnerability of rehearsal (Study III), and attention (Study IV) findings may provide a window onto shared cognitive processes relating to music and language within song. Although previous research studies have shown a decrease in short-term serial retention of visual-verbal items in the presence of irrelevant background sound (e.g., Colle & Welsh, 1976; Jones, 1993), this empirical series will extend this literature by exploring retrieval requiring familiar, known melody/lyrics production and lexical-semantic processing aided by LTM (Study I-II) in the presence of different melody/lyrics irrelevant background sound conditions, necessitating perceptual-motor processing for production/performance. Patterns of retrieval will then be compared to those from STM tasks of serial-digit recall (Study III) and non-serial missing-item identification (Study IV) not necessitating perceptual-motor processing, to help determine those auditory characteristics driving task recall accuracy.

CHAPTER II

AN OVERVIEW OF DISTRACTION BY TASK-IRRELEVANT SOUND

2.1 Introduction

Task-irrelevant extraneous sound is typically omnipresent, even though we only actively listen to a fraction of it. We cannot escape the constant exposure to auditory information that permeates our surroundings: Although we can ignore task-irrelevant visual material by closing our eyes, we have no "earlids" to enable us to ignore task-irrelevant auditory material. Extraneous noise emanates from various sources including traffic or office inhabitants/equipment (Hygge, Boman, & Enmarker, 2003; Schlittmeier & Hellbrück, 2009) and background music in shops or accompanying ceremonies, events, film or TV (Furnham, Gunter, & Peterson, 1994). The explosion of background noise pouring from our numerous and varied technical gadgets and work-spaces is unprecedented. Our ability to overcome distraction from some forms of task-irrelevant sound may improve with the development of top-down cognitive control over environmental stimulation (Joseph, Hughes, Sörqvist, & Marsh, 2018; Kattner & Ellermeier, 2020) or through routinely completing tasks under such conditions (Iwanaga & Ito, 2002; North, Hargreaves, & O'Neill, 2000; Schwartz & Fouts, 2003). We profess to be a sound conscious generation yet, we can work and talk through much of the sound, music or non-music, we "hear" without noticing. While we have become passive recipients of sound and may be unaware of its presence, the sound around us is nonetheless processed. One consequence of this passive processing of sound can be garnered from its detrimental impact on cognitive task performance (e.g., Colle & Welsh, 1976; Furnham & Strbac, 2002; Thompson, Schellenberg, & Letnic, 2011).

A significant body of research has shown a decrease in ability to retain short-term serial recall of visual-verbal items in the presence of background sound, termed the irrelevant sound effect (e.g., Beaman & Jones, 1997; Colle & Welsh, 1976; Jones, 1993; LeCompte, 1994, 1995, 1996; Salamé & Baddeley, 1982). The classical structuralist view (e.g., Baddeley, 1986) promotes the idea that this disruption is due to interference caused by the structural similarity between to-be-remembered and to-be-ignored items that cohabit a short-term store (Neath, 2000; Salamé & Baddeley, 1982). This chapter explores evidence that questions this viewpoint. One alternative explanation proposes task impairment from auditory distraction arises due to concurrent common processes operating during serial organisation of stimuli, identified as "interference by process" (Jones, Beaman, & Macken, 1996; Jones & Tremblay, 2000). Although much work in this area has related to serial recall (therefore, potentially paradigm bound) evidence for this interference-by-process view will be identified from serial short-term memory tasks that cast doubt as to the existence of the temporary storage hypotheses (e.g., Jones et al., 2004).

Chapter II comprises an outline of competing theories and key features pertaining to auditory distraction in the domain of short-term memory (STM) and semantic processing. Accounts of auditory distraction will be discussed first in the domain of serial recall, which will be evaluated based on characteristics of sound and tasks that determine disruption. Following this section there will be a focus on those aspects pertaining specifically to music, concluding with relevant aspects of semantic-auditory distraction.

2.2 The serial recall paradigm

Routinely, serial recall, one of the oldest experimental paradigms, has become the preferred medium for assessing storage and retrieval of verbal and non-verbal sequence patterns in STM. Almost all "benchmark" phenomena of STM such as the word-length effect (Baddeley, Thompson, & Buchanan, 1975), phonological similarity effect (PSE, Baddeley, 1966), and irrelevant sound effect (ISE, Colle & Welsh, 1976), that theories of working memory (WM, Baddeley & Hitch, 1974) should explain (Oberauer et al., 2018), have been identified using serial recall tasks. This paradigm requires recall of to-be-remembered information in the exact order in which it was initially presented. Recall can take place immediately after presentation of the last list-item or following a short retention interval (e.g., 5-10 second).

Originally termed the unattended speech effect, serial recall of visually presented letters was shown to be impaired by auditory stimuli comprising spoken text in a foreign language (Colle & Welsh, 1976). Later results that showed impairment to serial recall tasks produced by background nonsense speech (Salamé & Baddeley, 1982) defined the disruption as the irrelevant speech effect. Disruption on serial and free recall tasks with non-speech sounds e.g., tones (Beaman & Jones, 1997), by music and singing (Morris, Jones, & Quayle, 1989; Nittono, 1997), or differing tone sequences (Jones & Macken, 1993), and pitch-glides (Jones, Macken, & Murray, 1993), subsequently characterized it as the irrelevant sound effect (ISE). In this context the classical ISE describes difference in accuracy on serial recall tasks in the presence of irrelevant sounds that participants are instructed to ignore.

In order to remember items for serial recall, whether or not in the presence of task-irrelevant sound, attention needs to be directed towards the item. But those items also need to be rehearsed for later recall. According to some accounts (e.g., Jones & Macken, 1993; Salamé & Baddeley, 1982), the ability to complete this efficiently in the presence of irrelevant sound relates to the distraction effect—the ISE. Auditory distraction effects from serial recall studies are reliable (participants level of susceptibility at one time point is correlated with their susceptibility at another time point) and the ISE is generalisable, since it is shown to disrupt the majority of people (Ellermeier & Zimmer, 1997). Furthermore, serial recall underpins many everyday tasks such as mental arithmetic (Banbury & Berry, 1997, 1998) and memory for prose (Banbury & Berry, 1998). In addition, the ability to remember information in serial order in scholastic and workplace settings marks this method of determining any impact of irrelevant sound on task performance as justified and relevant (Gathercole & Baddeley, 1993).

***Acoustic masking*.**

To understand the notion that information read by the eye is affected by unattended information received by the ear, one needs to understand how the brain deals with such information. Since two distinct domains are utilized then disruption would seem to result from a breakdown in ability to direct attention towards the task requirement. Initially, Colle and Welsh (1976, p. 17) identified a phenomenon described as “acoustic masking in primary memory”. This idea stemmed from tasks that showed short-term serial recall of a series of visually presented letters impeded by the presence of spoken words in German, a language foreign to participants. Therefore, the effect was not driven by lexical processing. At first glance then, it seemed that some properties of speech could interfere with encoding/registration of visual to-be-remembered items despite them being in different modalities, potentially, due to some sensory confusion or masking (see Broadbent, 1982; Jones, 1994; Salamé & Baddeley, 1982). However, this view has been undermined by findings that irrelevant speech does not have to be presented simultaneously with to-be-remembered items to produce disruption. It can occur in a retention interval—wherein items should have previously gained unimpeded access to memory (Miles, Jones, & Madden, 1991). Moreover, placing to-be-ignored sound between visual to-be-remembered items creates as much disruption as during concurrent presentation, further undermining the masking view (Miles et al., 1991). Moreover, in studies involving two discrete sensory modalities the idea that some peripheral sensory masking occurs is inappropriate. Clearly then, another mechanism must be responsible for the ISE. Researchers subsequently endeavoured to discover factors influential and non-influential in determining the ISE.

2.3 Non-influential factors

Studies to assess effects of interference by irrelevant sound on non-auditory tasks have mainly utilized visual presentation of to-be-recalled verbal items. During their encoding and/or their rehearsal (retention period before recall) to-be-ignored background sound is played (e.g., Colle & Welsh, 1976; Salamé & Baddeley, 1982). The to-be-ignored sound's unique individual characteristics and research determining how and where the proposed disruption occurs, within processing, during perception, memory, or retrieval, have been fundamental to isolating its locus within the cognitive system. Whichever sound medium is employed (speech/non-speech), any level of fluctuation within an irrelevant sequence appears central to task impairment. Crucially, less task disruption is observed, especially when recall order is necessary, if the irrelevant sound sequence is steady (repeated) or has minimal acoustic variation (e.g., Jones, Alford, Bridges, Tremblay, & Macken, 1999). For example, irrelevant white noise[12] does not appear to cause disruption (Colle & Welsh, 1976). However, Tremblay and Jones (1999) identified broadband noise with centre-frequency changes does disrupt as does inserting quiet periods in a pitch glide. Therefore, the disruptive capability of an irrelevant sound is not so much reliant on its content but on changes to its acoustic properties (Jones, Saint-Aubin, & Tremblay, 1999).

Intensity.

Initially it was believed only very loud sounds disrupt recall task performance (e.g., Jerison, 1959). However, disruption of serial short-term memory by irrelevant sound is comparably affected irrespective of the irrelevant sound's intensity from 76dB (A) loud, to 48dB (A) whisper (Colle, 1980; Colle & Welsh, 1976; Ellermeier & Hellbrück, 1998; Tremblay & Jones, 1999). Tremblay and Jones (1999) also discovered that within serial recall level changes of sound within an irrelevant sequence, from 55dB (A) to 85dB (A) was not significantly disruptive, although changing, as compared to repeated, sequences of speech items were more disruptive.

Speech and meaning.

Irrelevant speech is not alone in its ability to cause disruption to cognitive focal tasks particularly serial recall (e.g., Jones, Madden, & Miles, 1992; Salamé & Baddeley, 1989). Non-speech sounds (e.g., Beaman, 2005), tones (e.g., Jones & Macken, 1993), instrumental music (e.g., Schlittmeier, Hellbruck, & Klatte, 2008), and environmental sounds (e.g., Perham, Hodgetts, & Banbury, 2013), can also disrupt performance.

[12] Random noise with flat power spectral density.

Given the lack of similarity between non-speech stimuli and to-be-remembered items in serial recall (usually digits or letters) this may be surprising to those assuming that disruption is produced via an interference-by-content at some level of similarity (e.g., phonological, semantic). However, important findings have undermined the role of speech and its meaning in serial recall experiments with auditory distraction. Spoken meaningful words alone are not sufficient or even necessary to create an ISE. For example, serial recall accuracy of nine digits visually presented during simultaneous presentation, through headphones, of 15 monosyllabic irrelevant words, such as CAT, PAD, CUP, and 15 irrelevant nonsense words, such as CAG, TAD, or GUP appeared equivalent (Salamé & Baddeley, 1982). Furthermore, other studies have shown that meaningless narrative speech (reversed speech or Welsh, an unknown language to the participants tested, Jones, Miles, & Page, 1990) are equipotent to meaningful speech in their disruptive capacity in the context of serial short-term memory tasks (e.g., Jones et al., 1990). Nonsense words and reversed speech have also been shown to be disruptively equivalent when compared to meaningful speech (Colle & Welsh, 1976; Jones et al., 1990; Marsh, Hughes, & Jones, 2009; but see LeCompte, Neely, & Wilson, 1997, for a limited effect that may, in fact, be due to acoustic complexity). Finally, between-semantic stream similarity is also ineffectual. For example, Buchner, Irmen, and Erdfelder (1996) found serial recall of two-digit numbers to be no more disrupted by different irrelevant two-digit numbers than by word combinations, phonologically related or not. Taken overall, these findings suggest that mere meaningfulness (the presence of semantic properties within the to-be-ignored sound) nor the semantic similarity between to-be-remembered and to-be-ignored items are influential characteristics in the context of the ISE (for a review, see Marsh & Jones, 2010).

Phonological similarity.

Originally, accuracy to serial recall tasks was thought to occur through disruption by the phonological content of to-be-ignored sound, as white noise bursts, (with no phonological properties) created hardly any effect to serial-digit recall (Salamé & Baddeley, 1982; see also Salamé & Baddeley, 1987). Furthermore, disruption attributable to the sound's phonological content was initially considered to be a function of the similarity in phonological content between to-be-remembered and to-be-ignored items (between-sequence phonological similarity). For example, Salamé and Baddeley (1982, Experiment 5) identified significant impairment to individual 1-9 serial-digit recall when participants were distracted by spoken irrelevant word sequences that contained the same phonemes (e.g., tun, gnu, tee) as random visually presented-for-recall digits (e.g., one, two, three). In addition, they both produced greater disruption than disyllabic words that were phonologically dissimilar (e.g., tennis, jelly, tipple (Salamé & Baddeley, 1982). However, the view that between-sequence phonological similarity is an important ingredient of the disruption of serial recall has been undermined by studies that have since identified that increasing between-sequence phonological similarity does not increase the disruption of serial recall task performance (Jones & Macken, 1995b; Marsh, Vachon, & Jones, 2008). For example, a sequence of syllables, such as "eff, kay, ell" in to-be-ignored words can disrupt a sequence of to-be-recalled words containing the same rhyming syllable such as "deaf, pay, bell" (Jones & Macken, 1995c). However, just as much disruption occurs to recall of a sequence of non-rhyming words such as "hat, cow, nest". As an irrelevant rhyming sequence (e.g., "sea, flea, key") was found to be less disruptive than dissimilar sequences (e.g., "hat, cow, nest" or "deaf, pay, bell") this points to disruption by acoustic changes between items within the to-be-ignored sequence rather than phonological properties (Jones & Macken, 1995c). Therefore, mere phonological content within the to-be-ignored sound, nor between-sequence phonological similarity, are influential contributors to the ISE (but see Eagan & Chein, 2012).

2.4 Structural accounts

2.4.1 Interference-by-content.

The idea that forgetting, rather than just resulting from a decay process through passing of time, could also be due to interference from similar memories that disrupt recall of target memories (Mueller & Pilzecker, 1900), has become established as a principle concept of interference within short-term memory. This view relates the degree of interference to similarity of traces between to-be-remembered and to-be-ignored items (Salamé & Baddeley, 1982). Forgetting, is typically deemed to result from new learning (retroactive interference) and previous learning (proactive interference). The hypothetical phonological short-term memory store (Salamé & Baddeley, 1982) is unable to cope with the simultaneous entry of both. In seminal work, McGeoch (1942) argued that both previous learning and new learning existed in memory in the form of memory traces. Impairment of original learning therefore, a direct consequence of the generated competition. According to structural interference-by-content accounts (Colle & Welsh, 1976; Neath, 2000; Salamé & Baddeley, 1982), the act of acquiring new information impairs memory for previous learning, the stronger competition with retrieval cues the greater impairment to recall. Consequently, forgetting is linked to an inability to deal with ensuing competition (McGeoch, 1942; Salamé & Baddeley, 1982). A number of structural accounts have been promoted to explain the classical ISE based on this idea of item identity.

2.4.2 Phonological store account.

The phonological store account (Salamé & Baddeley, 1982) developed as part of the WM model framework (Baddeley & Hitch, 1974) offers a structural interference-by-content account of the ISE. According to this account (for an overview see Baddeley, 1986), the mnemonic system for short-term retention of verbal material has two components, a passive phonological store, and an active articulatory control process. The phonological store is dedicated to temporary retention of verbal information, in the form of phonemic codes. Items within the store, represented by phonological codes, decay after a few seconds if not refreshed by the separate articulatory control process the operation of which maps onto sub-vocal rehearsal. Besides revivifying decaying phonological codes of to-be-remembered items, the articulatory control process is also responsible for a grapheme-to-phoneme conversion process whereby the written form of an item is converted into a phonological code for temporary storage within the phonological store. There are two modes of access into the phonological store. Auditory items have direct access. However, visual items must first be transformed into phonological codes (via a grapheme-to-phoneme conversion process) before being represented in the store. According to the model, an ISE is due to a clash between phonemes of to-be-remembered items—that achieve access to the phonological store through the grapheme-to-phoneme conversion process—and to-be-ignored items that gain direct access to the phonological store. Moreover, the extent of confusion was purportedly magnified by phonological similarity between to-be-remembered and to-be-ignored items (Salamé & Baddeley, 1982). Initial support for the phonological store account was derived from early experiments (Colle & Welsh, 1976; Jones et al., 1990) that showed inconsistent disruptive effects to recall by white noise despite being manipulated to act according to continuous speech (Salamé & Baddeley, 1987, 1989). This suggested that only irrelevant speech, not non-speech noise, had a direct and privileged access to the speech-specialised phonological store wherein it could interfere with phonological representations of to-be-remembered items (Gathercole & Baddeley, 1993; Salamé & Baddeley, 1982, 1989). However, the precise mechanisms by which phonological confusion between to-be-recalled and to-be-remembered items occurred within the store has not been detailed.

Compelling evidence has demonstrated that serial recall can be disrupted by task-irrelevant sound conditions other than speech (see Jones et al., 1992). For example, Salamé and Baddeley (1989) demonstrated that both instrumental and vocal music produced significant disruption of serial recall as compared with silence, while pink noise[13] was ineffective. That instrumental music produced disruption, albeit less than vocal music, suggested to the authors that an auditory filter existed at the interface of selective attention, and the phonological store was not completely efficient at barring non-speech sounds from entering the phonological store. Rather, Salamé and Baddeley (1989) proposed the existence of a filter that permitted sound material that was sufficiently speech-like to pass through the store, while it refused entry to non-speech sounds that were insufficiently speech-like, with pink noise being an example par excellence. Therefore, the authors held onto the assumption that the magnitude of the ISE was accentuated by increasing between-sequence phonological similarity, but they allowed scope for non-speech sounds to produce disruption to the rate that they are speech-like (Salamé & Baddeley, 1989).

Notions that between-sequence phonological similarity, and speech-likeness of irrelevant sound, determines the existence or magnitude of the ISE have been undermined by a raft of studies. For example, several studies have shown that increasing between-sequence phonological similarity has no impact on the magnitude of the ISE (Buchner et al., 1996; Jones & Macken, 1995b; Larson, Baddeley, & Andrade, 2000; LeCompte & Shaibe, 1997; but see Eagan & Chein, 2012). Despite suggestions that instrumental music conveys "speech-like" properties, which permit access into the phonological store—thereby allowing disruption from non-speech sounds—in terms of the WM model there is no accounting for an ISE produced by non-speech noise bursts of changing frequency (Tremblay, Macken, & Jones, 2001) and from non-speech-like sounds such as sine-tones (Jones & Macken, 1993). In addition, acoustic characteristics of sound, other than intensity, such as variability of timbre and pitch during the temporal presence of irrelevant material, dramatically impact on the manifestation of an ISE (the changing-state effect; Jones et al., 1992). This undermines any notion that the ISE is attributable to phonological content of the sound. Rather, it appears that the greater acoustic variability of irrelevant sound, the greater the disruption.

Primacy model.

[13] Random noise containing equal energy per octave.

Another account that questions dependency on phonological similarity between to-be-remembered and irrelevant auditory items in serial recall is the "primacy model" (Page & Norris, 1998). Here, order of items-to-be-recalled is crucial. Within this approach, irrelevant sound is capable of interfering not just with representations of to-be-remembered items, but also with their order (Norris, Baddeley, & Page, 2004). Following previous suggestion (Lewandowsky & Murdock, 1989), within the primacy model, item position is coded within a sequence, each item acting as a cue for the following item. According to this view, storing order information for successive list items decreases across list position to form a primary gradient. To recall list items in serial order, support is required from repeated operations to maintain rehearsal and develop associations between successive items within the sequence. However, a secondary activation for irrelevant auditory items removes activation from the primary order gradient containing the to-be-remembered items. Serial recall is reliant on the order in which information is stored relative to the strength of activation, which decreases according to list position as additional items are added. Following this argument, the idea that the ISE results from phonological similarity has now been overturned (Baddeley, 2007).

Feature model.

An alternative view postulating structural principles of interference-by-content is the "feature model" (Nairne, 1990; Neath, 2000). During a short-term memory task, two distinct features of relevant and irrelevant items are represented in a short-term memory store as two categories: modality-dependent (e.g., sensory, dependent on visual or auditory presentation), and modality-independent (e.g., semantic, or phonological identity). This model is characterised by the notion that processing of to-be-remembered items is impaired by irrelevant speech due to a feature-matching retrieval mechanism referred to as "feature adoption". Based on a +1 or -1 value, features of a to-be-remembered item are compared to a to-be-ignored item. If found to be the same the to-be-remembered item will be modified. The process of matching primary to secondary memory traces subsequently becomes degraded. According to the feature model, a loss of item leads to a loss of order information. This model explains the ISE as a consequence of feature adoption (Neath, 2000) whereby modality-independent features of irrelevant items overwrite corresponding features of to-be-remembered items. However, feature adoption has only been identified with speech sounds (Neath, 2000) pointing to a between-sequence similarity effect. In addition, the feature model does not account for findings that irrelevant non-speech disrupts serial recall performance in a similar manner to irrelevant speech (Jones & Macken, 1993; Tremblay, Nicholls, Alford, & Jones, 2000). Furthermore, it does not reconcile with the observation that acoustically varying irrelevant sound produces more disruption than non-varying sound—the changing-state effect (Jones et al., 1993). The assumption of the model is that a stream of changing-state, as compared to steady-state, items reduces the overall amount of attention applied to the focal task, resulting in a situation akin to divided attention. The changing-state irrelevant sounds therefore create an extra burden on focal task recall thereby giving rise to impaired performance (e.g., Neath & Surprenant, 2001). Research findings detailed in sections below, however, question the assumption that changing-state sound produces disruption because it draws attentional resources away from demanding focal tasks (e.g., Hughes & Jones, 2001; Hughes Vachon, & Jones, 2007).

2.5 Acoustic characteristics *versus* phonological content

Changing-state effect.

The mere presence of speech or sound does not determine disruption to serial recall (Jones et al., 1992; Jones & Macken, 1993). Of greater importance is the make-up—acoustic characteristics—of those irrelevant sounds. With the exception of intensity (Tremblay & Jones, 1999), acoustically variable irrelevant spoken letters, timbre, or pitch sequences, do have potential to create greater disruption to serial recall than less acoustically variable sequences of irrelevant sounds. The "changing-state hypothesis" (Jones, 1993)—which is part of the interference-by-process account (Jones & Tremblay, 2000) and perceptual-gestural view (Jones et al., 2004)—holds that changes in the acoustic stimulus, as would occur in changing speech letters (e.g., c, t, g, u, or varied pitch, A, C, F#, B), are more distracting relative to repeated sounds (e.g., c, c, c, c or constant pitch A, A, A, A). In this process, the acoustic make-up of the sound between utterances (differing syllables in contrast to single repeated syllables) is mainly responsible for disruption. Therefore, a series of changing-state sounds, speech or non-speech will invariably produce greater disruption of serial recall than a sequence of unvarying material. For example, staccato music, containing frequent changes of pitch/tempo is more disruptive than steady legato music (Klatte, Kilcher, & Hellbrück, 1995). Contrary to interference-by-content accounts (Neath, 2000; Salamé & Baddeley, 1982), but in accordance with the object-oriented episodic record (O-OER) model (Jones, 1993; see next section), speech and non-speech should be equipotent in their ability to disrupt serial recall performance as perceptual processing of those sounds will produce extraneous order information in a similar manner (Jones & Macken, 1993; Jones & Tremblay, 2000).

O-OER model.

The theory of an object-orientated episodic record (O-OER, Jones & Macken, 1995a), was a precursor for the interference-by-process account, now superseded by the perceptual-gestural approach (see below). This theory speculates the existence of a metaphorical “blackboard”, an episodic surface that retains auditory and visual material as objects in short-term memory. On this episodic surface order information between objects of auditory origin and visual origin is represented by pointers. Serial rehearsal via subvocalization is required for objects of visual origin to be represented on the blackboard and it is also responsible for maintaining the pointers between those objects thereby permitting their later retrieval. A second group of objects from auditory origin is represented automatically on the blackboard. According to the model, changing-state sounds are more disruptive than steady-state sound because each acoustic change between consecutive items yields an order cue, represented by a pointer on the episodic surface. Disruption occurs due to confusion between the episodic pointers representing the serial order of to-be-remembered items—generated deliberately via the process of sub-vocal rehearsal—and the pointers derived from perceptually registering the order of changes between auditory items. In the O-OER model, to-be-remembered items are therefore lost to recall due to loss of order cues, not item information (cf. Neath, 2000). A crucial difference between O-OER and the phonological loop hypothesis (Baddeley, 1986) is the focus on serial order (seriation; Jones, 1993).

2.5.1 Interference-by-process.

An alternative proposal, that critically questions whether interference results from item identity, investigates the relationship between processes used in the focal task and that automatically applied to the irrelevant sound (Jones, 1993; Jones & Tremblay, 2000). The initial impetus for the interference-by-process account came from observations that undermine the notion that similarity between to-be-remembered and to-be-ignored items causes increased forgetting, an interference-by-content view (e.g., Salamé & Baddeley, 1982). Critically, Jones and Macken (1993) found non-speech sounds such as tones were equipotent to speech in their potential to disrupt short-term memory performance despite conveying no phonemes or being speech-like. Of most interest, however, was that disruption of short-term memory produced by non-speech and speech sounds was jointly determined by acoustic and task factors. Jones and Macken (1993) reported that a single repeated item (a syllable or tone) produced much less disruption of serial recall performance than did a sequence of four different items (different syllables or tones that varied in pitch). However, this changing-state effect was not observed in the context of a missing-item task wherein participants are to report an item that is not presented from a well-known set (e.g., digits 0-9). Unlike for serial recall (Salamé & Baddeley, 1982), performance on the missing-item task is undiminished when opportunities for serial rehearsal are restricted by requesting participants repeatedly produce an irrelevant item (e.g., the word "the") aloud during encoding task material (articulatory suppression; Klapp, Marshburn, & Lester, 1983). This strongly suggests that the task, although making mnemonic demands, does not entail serial rehearsal. This very feature (that the task does not involve seriation), according to the interference-by-process view, gifts the task its immunity from disruption via the changing-state effect. Here then, it is argued that there is no interference by a similarity in process since the missing-item task does not typically involve seriation (see Hughes & Marsh, 2019) and thus there is no conflict between task process and a process applied pre-attentively to the task-irrelevant sound. In contrast, serial recall performance is susceptible to disruption via the changing-state sound precisely because it is underpinned by a seriation process (serial rehearsal). According to the interference-by-process account (Jones & Tremblay, 2000), this deliberate serial order processing of focal task material comes into conflict with a similar, but this time automatic and pre-attentive, process of serially organizing incoming irrelevant auditory items. Influential in this process is the passive auditory grouping or streaming corresponding to the changing-state of irrelevant items (Bregman, 1990).

Segregation and streaming.

According to the interference-by-process view (e.g., Jones & Tremblay, 2000) any disruption to focal task performance by irrelevant sound will be determined by factors related to streaming (Bregman, 1990). Segregation into separate auditory streams occurs when successive sounds within a sequence are compared to each other to determine their environmental origin. The interference-by-process view supposes that streaming occurs for distinct changes in successive sounds irrespective of whether the sound is attended (Macken, Tremblay, Houghton, Nicholls, & Jones, 2003). Therefore, conflict is generated between automatic seriation of irrelevant stimuli and deliberate serial order rehearsal of to-be-remembered items. As the focus is on acoustic properties rather than item identity this view contradicts the interference-by-content account. The interference-by process account makes the strong prediction that seriation processes must apply for both focal task and irrelevant sound processing, for disruption to take place. Moreover, an additional qualification is that such disruption will be dependent on the level of order-cue yielding acoustic variation (or changes-in-state) within task-irrelevant material.

Perceptual organisation.

Changing-state sound needs to be segmented into successively distinct elements to be disruptive on serial task performance (Macken, Phelps, & Jones, 2009). Hence humming has a limited disruptive effect because it provides little opportunity for segmentation as compared to sudden shifts of sound during singing or speaking (Morris & Jones, 1990). Perceptual organisation "looks" beneath the level of item presentation to identify their structure and formation for both auditory, in the formation of streams, and motor-output organisation, represented by sequences of sub-vocal gestures (Hughes & Jones, 2005; Jones, Hughes, & Macken, 2006, 2007). This perceptual-gestural view submits that incoming sequences are converted into gestural or articulatory forms. (The gestural element pertains to auditorily and visually presented lists, the perceptual element being mainly applicable to the auditory domain.) The magnitude of distraction is determined by the perceptual organisation of sound relative to its source—auditory scene analysis or streaming (Bregman, 2001).

Differences between successive sounds shown to produce stream segregation are highly distinct. For example, the order of sounds such as a burst of white noise, a tone, a vowel sound, and a buzz are harder to recall as compared to a less diverse sequence (e.g., burst of white noise, high and low pitched tones, and a buzz; Broadbent & Ladefoged, 1959; Warren & Obuzek, 1972). Individual sounds are segregated into separate streams if the degree of auditory change (e.g., pitch) exceeds a "fission" threshold, thereby reducing the distracting effect of changing-state (Jones et al., 1999). In addition, the point at which stream segregation begins is related to order information. For example, when zero to two to five semitones are introduced, a parallel increase of disruption to short-term memory serial recall is evident. However, at 10 semitones[14] disruption decreases, having reached the point at which higher and lower stream segregation occurs (Jones et al, 1999). For sequences in which tones of two different frequencies alternate, fission results in the perception of two streams within which little acoustic variation and thus order cues exist. This lack of order information results in less disruption to recall. In this view, the acoustic variation from opposing pitch-range

[14] One tone less than an octave.

musical instruments such as between flute and tuba should be less disruptive than sounds from a flute and oboe where acoustic differences are smaller.

Another example of how streaming is influential in modulating the ISE relates to the spatial location of irrelevant auditory stimuli, as sounds stemming from one location are grouped into a single stream. In other words, three-syllables looped to sound in both ears will create one stream and have enough distinction to fulfil the requirements of changing-state with subsequent order information leading to a pronounced ISE (Jones & Macken, 1995c; Jones et al., 1999). If delivery is modified, each sound presented to a different location at a rate to allow streaming (one syllable to right ear, one to left ear, one to both ears), disruption is reduced (Jones & Macken, 1995c) the syllables directed into three steady-state streams. As item identity is maintained for each delivery this finding does not fit with an interference-by-content view (Salamé & Baddeley, 1982), nor with the idea of attentional recruitment (Cowan, 1995, see section 2.5.2), as according to these approaches, item and location changes should create increased disruption. Irrelevant acoustic changes, therefore, need common ground (e.g., timbre) to cause marked disruption.

To clarify, if common ground is abolished, through acoustical change resulting from changes to a sound provider (e.g., different voice) the irrelevant sound will be more disruptive (Jones, Macken, & Harries, 1997). When items are produced by a common carrier (e.g., one voice) they form one stream hence order cues are strong. But when items occur from different voices, changes are too great to cohere into the same stream. Hence, separate streams conveying little order information are formed. A second or third voice added to the same spatial location results in a corresponding increase in memory disruption (Jones & Macken, 1995a). However, there is an optimum level at which changing-state is potent and this corresponds to when segmentation cues are clearest. When more than three voices are added to the same spatial location, disruption decreases. This decline of disruption in serial recall produced by many *vs.* few voices is termed the "babble effect" (Jones & Macken, 1995a; Jones et al., 1999). It is considered that the "babble" of multiple voices produces an auditory signal akin to a continuous amplitude that fails to provide a platform to enable segmentation, unlike the varied peaks and troughs from a single voice (Klatte & Hellbrück, 1993). Thus irrelevant order information—the ingredient driving disruption of serial recall—is impoverished with many *versus* few voices.

Impairment of serial recall by an acoustically less complex signal, sinewave speech (conveying less changing-state) was equivalent regardless of whether participants were aware the sinewave stimulus was speech (Tremblay et al., 2000). Continuous sinusoid produced from smooth oscillation of a steady tone (as opposed to more musically complex features such as attack, vibrato, or modulation) was fashioned through sinewave speech. Here, removal of phonetic structures resulted in sound resembling computer beeps or whistling (Sheffert, Pisoni, Fellows, & Remez, 2002). However, natural as compared to sinewave speech, produced greater impairment reinforcing less complex sound, with less changing-state, results in less task impairment. Similarly, disruptive effects of speech played forwards or backwards are equipotent (LeCompte et al., 1997). However, if spoken items are limited (e.g., four items) speech played backwards is less disruptive (LeCompte et al., 1997), thus continuity of continuous speech (e.g., Jones et al., 1990) appears to equipoise auditory complexity and subsequent segmentation cues. It could be argued here, however, that the nature of the words used in the only study demonstrating additional disruption from normal as compared to reverse speech (LeCompte et al., 1997), "hey, you, me, no" were attentionally-capturing (see 2.5.2) which would account for them being more potent disrupters than their reversed counterparts. Although the changing-state hypothesis identifies the nature of irrelevant stimuli that produce disruption on serial short-term memory, it does not identify the specifics of the required mechanism.

Task process specificity.

Building on initial observations by Jones and Macken (1993) subsequent research has offered further support for the interference-by-process account through manipulating task-requirements. For a pronounced changing-state ISE to be routinely observed, seriation within a focal task is necessary and is the main determinant of disruption (Beaman & Jones, 1997, 1998; Jones & Macken, 1993; Jones & Tremblay, 2000). Therefore, according to the interference-by-process account, all tasks involving seriation between target items and irrelevant stimuli should be susceptible to disruption including tasks that make similar cognitive demands but differ in requirement for completion. For example, a probe-recall task (requesting a specific item in relation to another within a to-be-remembered list) necessitates order information and is vulnerable to disruption, but a missing-item task (identification of one item missing from a to-be-remembered list) can be accomplished without necessitating seriation and appears invulnerable. Although at first glance free recall (learned items recalled in any order) does not require serial order rehearsal, an ISE has been evidenced for free recall of lists of unrelated words or consonants frequently triggered by participants use of seriation (e.g., Beaman & Jones, 1998; LeCompte, 1994). In addition, a changing-state ISE does not require verbal information, as memory for correct series of dots presented in different spatial locations on a computer screen also demonstrate a changing-state effect (Jones, Farrand, Stuart, & Morris, 1995). This suggests that amodal order mechanisms are also susceptible to disruption via an interference-by-process.

The perceptual-gestural account.

The concept of interference-by-process is encompassed within an emerging account of short-term memory phenomena that has been coined the perceptual-gestural account (e.g., Jones et al., 2006), or perceptual-motor view (Hughes & Marsh, 2017). This view assumes that short-term memory performance is parasitic on processes and systems that are general-purpose mechanisms not specifically dedicated to memory. These general-purpose systems include deliberate processes such as vocal actions that are involved in producing coherent sequences of actions (e.g., Rosenbaum, 2009) and involuntary pre-attentive, perceptual processes that perceptually organize incoming sequences of sound (Bregman, 1990). The motor component of the perceptual-motor view refers to a motion-action emulator that runs in an emulation mode. This output-planning device operates in the absence of overt actions, generating a "forward model" of an action. This forward model comprises instructions to effectors and the sensory consequence of commands enables comparison of planning with the intended action as well as its fluent production (e.g., Grush, 2004; Schubotz, 2007). The perceptual-motor view has been offered as an account for a number of short-term memory phenomena (see Hughes, Marsh, & Jones, 2011; Jones et al., 2004, 2006; Macken, Taylor, Kozlov, Hughes, & Jones, 2016) including the irrelevant sound effect.

The perceptual-motor view eschews the notion that inner speech (sub-vocal rehearsal) serves the function of refreshing decaying memory traces within a mnemonic store or space (e.g., Baddeley, 2007). Rather, the skill of speaking (overt or inner-speech) is exploited to bind together items—that have no pre-existing syntactical or grammatical cues as to their order—into a single temporally-extended motor-plan on a common carrier. Speech, being necessarily sequential (due to the possession of only one vocal tract and the emulation of movement therein) is thus well suited to embodying the serial order of items and co-articulatory and prosodic characteristics of natural speech can further imbue the motor-plan with serial order cues (Macken, Taylor, & Jones, 2014; Woodward, Macken, & Jones, 2008). The perceptual-motor's explanation of the irrelevant sound effect stems from a proposed under-specification of action-parameters problem (see Neumann, 1987, 1996). Speech as a skill is an abstract entity, it is guided by generic sets of action parameters that govern movements of the vocal tract to produce coherent speech. However, this general set of action parameters is initially unpopulated by specific to-be-remembered content (phrases, words, phonemes etc). As a consequence of this under-specification, any prospective irrelevant input, even if not specifying action-parameters appropriate to the task, can threaten to populate the motor-plan thereby producing interference. Therefore, in the context of serial recall, limitations in performance on the perceptual-motor view, are not due to limitation of mnemonic capacity within a static short-term memory store or space that is determined by the existence of items prone to decay, interference, or displacement as interference-by-content accounts hold (Baddeley, 1986; Neath, 2000).

The perceptual-motor view explains the greater disruption of serial recall observed from changing- as compared to steady-state sequences as being due to the motor-plan's susceptibility to the obligatory process involved in perceptually organising sound into streams (Hughes & Jones, 2005; Jones & Macken, 1993). Providing successive sounds within a changing-state sequence are acoustically similar to one another (in terms of frequency or timbre) and share a common ground (e.g., voice, musical instrument) they are assigned to the same stream and their serial order sequence formed via the assignment of successive cues to the same stream results in it being a candidate for populating the motor-plan. Disruption to serial recall produced by changing-state sound thus represents a competition-for-action. Because auditory items do not require to be verbal to be streamed, the perceptual-motor view explains why changing sequences of non-speech sounds such as non-vocal music (Schlittmeier et al., 2008) and tones (Jones & Macken, 1993) can give rise to disruption.

2.5.2 Attentional capture: Deviation effect.

While interference between two contemporaneous serial-order processes has been put forward as its explanation for the changing-state effect, the interference-by-process view asserts that the changing-state effect is a qualitatively distinct effect to that produced by attentional capture (Hughes, Vachon & Jones, 2007). On the interference-by-process view, unexpected sounds within an irrelevant auditory stream capture attention while changes between successive sounds are processed pre-attentively. Identified as "deviants" or "oddballs", unexpected sounds capture/divert attention from a focal task irrespective of the processing involved in that task (Hughes et al., 2007; Vachon, Labonté, & Marsh, 2017). Disruption produced by deviant sounds to focal task performance has been coined the "deviation effect" and is observed in serial and non-serial tasks such as the missing item task (Hughes et al., 2007). Unlike the susceptibility of the deliberate seriation process to the changing-state effect, it would appear that many non-seriation-based processes are susceptible to attentional capture. Attentional capture by task irrelevant sounds has been dichotomized into specific attentional capture—produced by a high valence word (e.g., Marsh et al., 2018), taboo word (Röer, Körner, Buchner, & Bell, 2017) or one's own name (Röer, Bell, & Buchner, 2013)—or aspecific attentional capture as produced by a sudden acoustic change within an auditory sequence (e.g., A, A, A, **D**, A).

Specific attentional capture occurs when the sound's content is responsible for gifting the sound its power to divert attention. An example of this form of attentional diversion is "the cocktail party effect", wherein an individual's ability to retain focus on one source of information within an environment is hindered when one's own name as compared to a control name, is spoken within a task-irrelevant source of information (Cherry, 1953; Moray, 1959). Information personally significant to an individual such as their own name or sounds related to an individual's current state (e.g., the sound of frying for a hungry person), captures attention regardless of the context in which they occur. For example, Röer et al. (2013) found more errors were made in subsequent serial recall, when participants' own names were presented as part of irrelevant sound thereby supporting previous literature in the context of dichotic listening (Moray, 1959; Wood & Cowan, 1995). In contrast, aspecific attentional capture is context-dependent, occurring when a sound captures attention because of the context in which it occurs such as the sudden onset of speech following a period of quiet (see Eimer, Nattkemper, Schröger, & Prinz, 1996) or an auditory token that violates the prevailing pattern of stimuli.

Cognitive control.

Cognitive control has selective effects on the disruption produced by changing-state sounds (changing-state effect) and deviant sounds (deviation effect). The role of top-down cognitive control in modulating the deviation effect but not the changing-state effect has been observed within studies wherein top-down control is addressed through contextual factors (e.g., high task-difficulty) or stable dispositions for attentional control. In relation to top-down cognitive control, there are currently three lines of evidence supporting the notion that the deviation effect and changing-state effect are qualitatively distinct forms of auditory distraction

First, when participants are forewarned about the nature of an impending auditory stimulation (e.g., "deviant", "changing-state"), the disruption that a deviant produces is eliminated, but the disruption produced by a changing-sequence of sound is undiminished (Hughes, Hurlstone, Marsh, Vachon, & Jones, 2013). Forewarning allows an opportunity to mentally fashion a neural model about the nature of the upcoming sequence that provides individuals with an expectation about a forthcoming deviant and reduces its capacity to capture attention from a focal task. Second, when a stimulus is harder to read, the deviation effect is reduced but the changing-state effect is not affected (e.g., Hughes et al., 2013; see also Marsh, Campbell, Vachon, Taylor, & Hughes, 2019). Increased task difficulty potentially increases engagement with the task perhaps potentiating a blocking mechanism for the call to attention by the auditory deviant (Hughes et al., 2013; Marsh et al., 2019). Third, individual differences in attentional control as measured with tasks measuring working memory capacity (WMC, e.g., OSPAN[15]) have shown reliable correlation with the deviation effect but not the changing-state effect: Specifically, higher working memory was correlated with a smaller deviation effect. However, it should be noted that this negative correlation between WMC and the deviation effect was not found by Körner, Röer, Buchner, and Bell (2017). They found no evidence of any relationship between WMC and the deviant or changing-state effect. However, conclusions for this study are somewhat undermined by the generally small number of participants manifesting a deviation effect that potentially occludes the discovery of a correlation.

To summarize, in contrast to the changing-state effect, the deviant effect is tempered by factors related to top-down cognitive control including foreknowledge (Hughes et al., 2013), task-difficulty (Hughes et al., 2013; Marsh et al., 2019) and WMC (Hughes et al., 2013, Marsh, Vachon, & Sörqvist, 2017; but see Körner et al., 2017). The deviation effect, therefore, would appear to be a phenomenon for which the interference-by-process account alone does not work but rather points to the requirement for a duplex-mechanism account.

2.5.3 Duplex-mechanism account.

The duplex-mechanism account promotes two discrete forms of auditory distraction (Hughes et al., 2007). The first, interference-by-process mechanism relates to a processing overlap, and subsequent conflict, between the pre-attentive processing of the serial order of changes within an acoustically varying auditory sequence and the deliberate seriation process applied to focal task materials in service of their retention (Hughes, Tremblay, & Jones, 2005). The second mechanism relates to attentional capture and is manifest in the deviation effect (Hughes, Vachon, & Jones, 2005; Hughes et al., 2007). In this view, the changing-state effect and the deviation effect produce

[15] OSPAN - operation span.

disruption to task performance via two different mechanisms. The changing-state effect is produced by an interference between two seriation based processes and the deviation effect is attributable to attentional capture (Hughes et al., 2005; Hughes, 2014; Hughes et al., 2013). The task-specificity of the changing-state effect and the task-specificity of the deviation effect (Hughes et al., 2007) supports the duplex mechanism account as does the influence of top-down cognitive control over specific (e.g., distracter valence; Marsh et al., 2018) and aspecific (e.g., sequence violation; Hughes et al., 2013; Marsh et al., 2017) attentional capture effects. Further, the reduced disruption over trials (habituation) observed with the auditory deviation effect (Vachon, Hughes, & Jones, 2012), but not the changing-state effect (Hughes et al., 2005), buttresses the claims that the two effects are underpinned by different mechanisms.

Unitary account.

Despite accumulating evidence for the duplex mechanism account (Hughes, 2014; Hughes et al., 2013), some scholars retain theoretical ideas on auditory distraction that align with a unitary account of auditory distraction (Cowan, 1995; Elliott, 2002; Röer et al., 2013, 2015). The unitary account is derived from the broader *embedded processes model* (Cowan, 1995, 1999, 2001; Cowan et al., 2005). In this view, all auditory distraction effects are attributable to exogenous attentional capture. The account asserts that the changing-state effects arise because each discrete change within a sound stream captures attention (orienting response; OR, Sokolov, 1963) away from to-be-remembered items in a capacity-limited focus of attention. This results in loss of accessibility to those items and therefore disrupts short-term memory performance. In contrast to a changing-state sequence, the attentional capture response to a steady-state sequence rapidly habituates. A widely held assumption is that features of incoming auditory stimulation are incorporated into a mental description called a neural model (Sokolov, 1963). This acts as a template against which to compare incoming stimuli. If there is a mismatch between an incoming stimuli and the representation of that stimuli within the neural model, an orienting response (attentional capture) is elicited. With repeating stimuli (e.g., D, D, D, D) a neural model is easily fashioned and an OR quickly habituates. However, for changing-state stimuli (e.g., D, K, R, L) a neural model is more difficult to specify, and every letter-to-letter change within the stimulus will capture attention. Thus, the unitary model explains the changing-state effect as resulting from the repetitive capture of attention induced by each change within the sequence (Hughes, 2014).

A number of findings undermine the explanation of the changing-state effect offered by the unitary account. For example, the account supposes that the impact of a single deviant should have more attentional capturing power in a steady-state than a changing-state sequence. This is because each change within a changing-state sequence should already be attention capturing. However, the impact of a single deviant item (a change in voice) has been shown to be no less potent at capturing attention when in the context of a changing, as compared to steady-state irrelevant sequences (Hughes et al., 2005). Further, if each change within a sequence captures attention then it follows that repetitions of the same item within that sequence should give rise to habituation of the orienting response for that item. For example, an orienting response to the letter C in a sequence "CACACACAC" would be expected to habituate more rapidly than to the letter C in the sequence "CABCABCAB". However, the evidence for this is mixed at best (Tremblay & Jones, 1998; see Bell, Röer, Lang, & Buchner, 2019). Although in contrast to the aspecific attentional capture effects (e.g., as produced by one's own name) are reduced following repetition suggesting that habituation of the orienting response produced by those stimuli occurs (Röer et al., 2013).

Despite success in accounting for data of the classical ISE, such as effects of non-speech sounds (Jones & Macken, 1993), the changing-state effect (Jones, 1993), auditory streaming (Macken, Mosdell, & Jones, 1999), perceptual organisation (Macken et al., 2003), and seriation (Beaman & Jones, 1997), one potential weakness of the interference-by-process framework is its reliance on serial recall. In order to counter the assumption that the interference-by-process concept is paradigm-bound it needs to be explored in other memory domains. This thesis, in addition to serial recall, will explore music and semantic auditory distraction during melody and lyrics retrieval and production of familiar song. Therefore, the following section details music-auditory distraction in relation to the irrelevant sound paradigm which will be followed by a section related to semantic-auditory distraction (2.10).

2.6 Music-auditory distraction

A number of studies suggest that the cognitive system appears to involve neural models to fashion prediction or expectation about upcoming notes from previous notes within a musical phrase. McCullough Campbell and Margulis (2015), for example, comment that one musical note within a melody carries implications, in terms of pitch or rhythm, for the next note within the phrase. The way music is encoded leads the system from the first note involuntarily through to phrase end or final cadence.

Described as a "cognitive itch" (Kellaris, 2001, p. 66) the cognitive system processes the music it is hearing, be that during active listening or from passive exposure. This automatic processing of sound is serial-order based and it is possible that melody (in particular familiar melody), may yield order cues *above and beyond* those pre-attentively processed from non-musical irrelevant sound as part of the auditory streaming process. Irrelevant music, whether heard involuntarily, or deliberately, is inescapable, and therefore conveys the potential to unavoidably reduce cognitive task performance (Iwanaga & Ito, 2002; Pring & Walker, 1994).

2.6.1 Music and the irrelevant sound effect (ISE).

A growing number of studies have investigated whether short-term maintenance of musical sequences is sub-served by the same mechanisms, or a mechanism different, to that responsible for the immediate retention of verbal material. This work assumes a structuralist approach and seeks to determine whether the so-called phonological loop deals with retention of music, as well as linguistic, content, or whether an independent system or store deals with music content. Within the domain of the WM model, traditionally it had been assumed that non-verbal information required verbal recoding to gain access to the phonological store. However, the possibility that music gains access to the phonological loop has been raised (Baddeley, 2012; Salamé & Baddeley, 1989) as has the possibility that verbal and music information tap at least partially overlapping cognitive resources (Atherton et al., 2018; Williamson, Baddeley, & Hitch, 2010a). However, at odds with the notion that the phonological loop can encode and process music information, is an opposing account that supports the further fractionation of WM to include a specialised short-term memory store for music (Berz, 1995; Pechmann & Mohr, 1992).

Cross-modal interference paradigms.

Evaluation of the underpinning of music retention typically makes use of cross-modal interference paradigms (Atherton et al., 2018; Williamson et al., 2010). Within this work "modal" or "modality" refers to music against linguistic information, rather than presentation modality (e.g., auditory *vs.* visual). Within the cross-modal auditory interference paradigm, participants are presented with to-be-remembered stimuli (non-music-verbal or music [e.g., tonal information]), followed by potentially interfering information (non-music-verbal or music) before attempting to recall to-be-remembered information (Atherton et al., 2018; Pechmann & Mohr, 1992). The level of interference that occurs for each type of to-be-remembered stimulus—non-music verbal *versus* music—is used to infer the nature of the cognitive system that sub-serves immediate retention of verbal or music information (cf. Williamson et al, 2010). For example, if verbal and non-verbal music information are processed within independent systems or stores, then it should be possible to observe modality-specific interference (e.g., irrelevant sound effect). If, however, verbal and non-verbal music information are processed by the same system then interference observed should not be specific to a modality (irrelevant music should interfere with verbal retention as much as it does with music retention, and vice versa).

Using the cross-modal interference paradigm, Williamson, Mitchell, Hitch, and Baddeley (2010b) investigated participants' ability to recognize musical tone sequences visually-displayed as musical notation in a later same-different auditory recognition test. It is generally considered that effective performance on visual-to-auditory comparison tasks requires participants to undertake "notational audiation" or "musical imagery" for the visual notation, that itself requires a degree of musical literacy. Williamson et al. (2010b) showed that auditory recognition was more greatly impaired by irrelevant tones than speech (see also LeCompte et al., 1997; Pechmann & Mohr, 1992; Schendel & Palmer, 2007). Further, this visual-to-auditory method also showed participant's ability to recognize letter sequences was disrupted more by irrelevant spoken digits, than tones. The authors propose this double-dissociation indicated that different processes underpin the retention of musical and verbal information. However, the generalisability of these findings are questionable: Given that musical expertise is required to encode the visual notation in the case of tones, it is unknown whether the same pattern of results could be found for individuals without musical knowledge. A feature of the results from Williamson et al. (2010b) that cohered with previous work is that performance for verbal recognition was superior to tonal recognition (Schendel & Palmer, 2007). They concluded that this could in part be due to limited memory practice for musical notation (variable according to participant's musicianship) as compared to visual-verbal material.

Further investigation of similarity or disparity of processing verbal and non-verbal material using the visual-to-auditory method have deployed conditions wherein covert vocal-motor processing is prevented. In this technique, participants are presented with tones in notation form, with a task of comparing them with an auditorily presented tone sequence using a same-different judgement. This visual-auditory variant is also typically compared with performance wherein the tone sequence is auditorily presented (auditory-auditory; Schendel & Palmer, 2007). Schendel and Palmer (2007) report that preventing rehearsal by requesting participants to sing "la" or articulate the word "the", led to equivalently poorer recognition performance on visual-to-auditory and auditory-to-auditory versions of the tone matching task, as compared with when rehearsal was not prevented via suppression. Similar results were reported for a condition in which digits were presented auditorily or visually, however, digit sequences were somewhat better recognized. The typical advantage of auditory presentation over visual presentation under suppression conditions (Jones et al., 2004) was not observed for notes, which has several implications. For example, visual notations may not be converted into an auditory imagery representation via an articulatory process, they may gain direct access to auditory imagery or they are represented differently (e.g., visuo-spatially) in the cognitive system. However, Schendel and Palmer (2007) provided some evidence against this latter possibility.

On the whole results of Schendel and Palmer (2007) suggest that similar covert rehearsal mechanisms (e.g., seriation via motor-planning) may play a role in short-term remembrance of music and verbal material. Although there is some inconsistency between verbal and note retention under suppression conditions, it is plausible that this reflects a change in strategy-use, rather than evidence of a domain-specific storage for music (cf. Berz, 1995). The general advantage observed for recognising digits over notes may be the product of skills or practice developed from the more frequent scenarios in which numbers are memorized over tones (Schendel & Palmer, 2007) or differences in encoding (Williamson et al., 2010b). However, findings from interference paradigms are mixed which suggests that assertions as to the degree of similarity in the way verbal and music sounds are processed for auditory short-term memory should be made with appropriate caution (Williamson et al., 2010b).

It is possible that cross-modal auditory interference studies fail to reveal clear evidence for specificity of short term memory for music due to lack of accounting for musical expertise. Schulze, Zysset, Mueller, Friederici, and Koelsch (2011) showed that neural activation related to working memory for linguistic and musical information included a greater neural separation for verbal and musical information for musicians as compared with non-musicians. Perhaps the clearest evidence for a verbal *versus* music distinction comes from tasks in which to-be-attended material requires deciphering of notation. However, only individuals with musical expertise can achieve this decoding. Therefore, it remains to be seen whether musicians as opposed to non-musicians demonstrate evidence for independence of processing music *versus* verbal material within WM.

It should be noted here that in the typical irrelevant sound paradigm, to-be-remembered material is presented visually and to-be-ignored material is auditorily presented. Although uncommon, however, auditory presentation of memoranda and irrelevant material has also been studied in an irrelevant sound paradigm (Nicholls & Jones, 2002; Schlittmeier et al., 2008). For example Schlittmeier et al. (2008) examined the impact of irrelevant music on one group of participants who *read* to-be-remembered digits, with a second group that *heard* the digits. A concurrent to-be-ignored instrumental baroque staccato passage (detached articulation) produced greater disruption than a to-be-ignored legato passage (smooth meditation) regardless of the modality of presentation. Since staccato music conveys more changing-state information than legato music, results support similar changing-state effects with auditory as with visual presentation of to-be-remembered items. The authors rule out alternative explanations of the apparent changing-state effect with auditory to-be-remembered presentation by incorporating a condition in which the irrelevant music started one second after presentation of to-be-recalled items. This negates alternative explanations, based on acoustic confusion or masking, impaired encoding of to-be-recalled digits and listening effort (Schlittmeier et al., 2008). Thus, the irrelevant sound effect observed with auditory presentation of to-be-remembered and to-be-ignored material appears amenable to the same mechanisms as those involved in its cross-modal counterpart (e.g., Nicholls & Jones, 2002). Further, the fact that tones and instrumental music produce qualitatively, if not quantitatively, similar disruption to visual-verbal serial recall (e.g., Jones & Macken, 1993, Klatte & Hellbrück, 1993; Salamé & Baddeley, 1989) suggests that speech and non-speech sound are subject to similar perceptual processing within the cognitive system. That is, the pre-attentive processing of changes within an unfolding auditory sequence (e.g., revealed by pitch and rhythm) has a similar consequence for retention of to-be-remembered material. However, whether to-be-recalled sequences of verbal material (e.g., speech) and non-verbal material, music material (e.g., tones) are processed similarly via covert, vocal-motor processes, or by different rehearsal/retention systems is currently unknown but is the subject of intense inquiry (Defilippi, Garcia, & Galera, 2019).

2.6.2 Pitch *versus* rhythm.

The question of whether different features of sound, usually intertwined within musical excerpts such as melody, can have independent effects on cognitive performance have been addressed in a number of studies (Furnham & Strbac, 2002; Hallam, Price, & Katsarou, 2002; Pring & Walker, 1994; Schellenberg, 2005; Thompson, Schellenberg, & Husain, 2001). Two such features are pitch and rhythm and the isolation of these two fundamental elements to determine their impact on performance has centred mainly on music's use as a mnemonic (e.g., Rainey & Larson, 2002). Attempts to discriminate possible independent effects of these musical elements have revealed separable effects on serial recall. Silverman (2007) found serial digit recall performance to be better when the items were paired with rhythm than when they were paired with speech, pitch, or an unfamiliar melody (pitch combined with rhythm) conditions (see also Silverman, 2010). This supports the notion that rhythmic patterns can assist chunking digits into memorable recall units (Prickett, 1974; Schellenberg & Moore, 1985). However the manipulations deployed by Silverman (2007, 2010) cast the sound and musical elements as part of the same object (e.g., via a female alto voice) that participants may opportunistically capitalise on. Therefore the paradigm has little relevance to when musical material is strictly irrelevant to the focal task as in the case with the classical irrelevant sound paradigm.

2.6.3 Vocal-music-auditory distraction.

Within the context of the classical irrelevant sound paradigm, lyrics, when accompanying melody, invariably result in greater disruption to serial recall performance than melody alone (e.g., Alley & Greene, 2008; Nittono, 1997; Perham & Vizard, 2011; Salamé & Baddeley, 1989). Furthermore, Nittono (1997) suggests familiarity with the non-speech musical excerpt may be important in governing disruption. Nittono (1997) found that compared to quiet, familiar instrumental music (played forwards) disrupts serial recall. However, unfamiliar instrumental music (played backwards) did not disrupt performance relative to quiet and failed to disrupt performance compared with familiar music.

Intriguingly, un-vocalized nursery rhymes produced more disruption than classical instrumental music and quiet (Pring & Walker, 1994). Although they failed to demonstrate any priming of associated lyrics in a word-stem completion task. Pring and Walker (1994) concluded that hearing the overlearned un-vocalized nursery rhymes resulted in automatic activation of lyrics within the phonological loop of the WM system. Herein, their representation produced confusion with that of the to-be-recalled items. Thus, an irrelevant speech effect obtained from "imagined speech". That the mere presence of melody can automatically activate lyrics (and vice-versa) is well known (Peretz & Coltheart, 2003; see Chapter I, 1.3.1, and Chapter IV, 4.1.2). However, the basis of its disruptive effect in Pring and Walker's (1994) study could be debated. Rather than assume that disruption arises due to an interference-by-content within the putative phonological store (Pring & Walker, 1994), it is possible to argue that it reflects an interference-by-process. According to Zatorre and Halpern (2005) the ability to "hear" aspects of music without actual sound, utilizing the "inner ear" (Zatorre & Halpern, 2005), has been described as a form of subvocalization (see Chapter I, 1.7). Therefore, automatic retrieval of lyrics could compete with to-be-recalled items for the vocal-motor planning system. An alternative, but not necessarily conflicting, account could be that representation of lyrics within WM might require a process of inhibition to remove them from the system or motor-plan. Disruption of task performance may therefore represent an overhead of this inhibitory process (cf. Marsh et al., 2018). While these explanatory views may appear appealing, there are aspects of Pring and Walker's (1994) work that require consideration. Some necessary conditions were not included to fully determine their variability. For example, Pring and Walker (1994) did not use irrelevant music with lyrics, speech, or an unfamiliar music comparison. This means that a number of other features within the irrelevant material, such as the musical elements of harmony, tempo, and dynamic, rather than implied linguistic properties could determine the additional disruption produced by the unvocalized nursery rhyme condition against the classical instrumental condition (cf. Klatte et al., 1995).

2.7 Arousal, mood, and selective attention

Another possible explanation for the modulation of task performance produced by irrelevant sound is due to the effect music has on arousal and mood. The arousal-mood hypothesis (Thompson et al., 2001) has been offered as an explanation for the benefit derived from listening to music by Mozart prior to undertaking a focal task (typically visuo-spatial), the so called Mozart effect (Rauscher, Shaw, & Ky, 1993, 1995). Although difficult to replicate (Pietschnig, Voracek, & Formann, 2010) the effect can generalises to different genres (e.g., "The Blur effect", Schellenberg & Hallam, 2005), or narration of a Stephen King novel (Nantais & Schellenberg 1999). Novel music can also affect arousal, mood, and attention (Bosch, Salimpoor, & Zatorre, 2013). However differences in participants' arousal threshold, personality, and emotional response need consideration (see extrovert *vs.* introvert experiments; Cassidy & MacDonald, 2007; Dobbs, Furnham, & McClelland, 2010; Furham & Strbac, 2002; North & Hargreaves, 1997). Since the benefit of increased arousal or positive mood to a cognition task occurs from previous exposure to a sound, it is uncertain whether any gain should be seen in performance when the sound accompanies task performance. In the context of serial recall, several studies suggest this would appear unlikely (e.g., Perham & Vizard, 2011; Schlittmeier & Hellbrück, 2009; for music disruption to focal task despite it including positive mood, see Threadgold, Marsh, McLatchie, & Ball, 2019).

2.7.1 Familiar *versus* unfamiliar music.

Memory for familiar melodies, but not necessarily the temporal process of learning, is generally understood to denote musical semantic memory. However, there are limited studies wherein semantic memory is tapped by sound of a familiar melody. Salamé and Baddeley (1989) determined that a familiar musical style or genre (operatic arias sung in French or Italian, and either nineteenth century or modern popular instrumental music) had no effect on serial recall accuracy. However, the notion that familiar, as compared to unfamiliar, irrelevant background music may draw attention away from a target task (e.g., McCorkell, 2012), is of particular interest in relation to the current thesis. Very few studies have manipulated familiarity within irrelevant sound. At some level of granularity, familiarity has been manipulated in studies comparing normal against reversed speech on serial recall performance, but familiarity here relates to the phonetic and acoustic properties of the sound. Moreover, the fact that, regardless of its direction, speech is equally disruptive of serial recall (Jones et al., 1990), suggests that familiarity in this sense, does not determine disruption. However, if one considers familiarity to be based on level of exposure and personal significance, then the own-name effect (e.g., Röer et al., 2013) might be considered an index of disruption produced by familiarity.

Within the domain of music, however, limited studies have manipulated familiarity with a to-be-ignored musical excerpt and the evidence from those studies is mixed. For example, in unpublished work, McCorkell (2012) found that a familiar song was more disruptive than an unfamiliar song even when the song material was broken down into conditions wherein only the instrumental, spoken-lyrics, and spoken-lyrics over instrumental was presented. This work suggests that properties other than its acoustic variation (e.g., its familiarity) drive the disruptive impact of an irrelevant sound stimulus. However, caution should be exercised here since a number of interpretations of these results are possible. For example, since words were used as to-be-recalled stimuli, it is possible that some semantic interference could ensue from a clash between the obligatory processing of the semantic properties of the irrelevant material (e.g., via the activation of lyrics in the instrumental condition) and the deliberate semantic processing of the to-be-recalled words (Marsh & Jones, 2010).

Earlier findings by Perham and Vizard (2011) appear contradictory to those of McCorkell, 2012. Perham and Vizard presented participants with music chosen by the participants themselves (assumed familiar song) and music that the experimenter chose (assumed unfamiliar song). Despite the variety of genres and diverse musical elements within the participants' self-chosen music, serial recall of memorized consonants was no different in the presence of participants' self-selected music than in that chosen by the experimenters and a changing-state condition (random permutations of the digits 1-9). Therefore, familiar music (chosen by the participant) was no more disruptive than unfamiliar music and participants' preference (they liked their self-chosen music while disliking the experimenter-selected music) was also an uninfluential factor. One problem with the inference that musical excerpt familiarity plays no role in its disruption of serial recall is that the experimenter-selected, unfamiliar, music ("Thrashers", a fast-raucous metal genre) may have had less changing-state properties than the participant's chosen music and thus the unfamiliar and familiar music conditions are poorly matched in terms of their acoustic qualities. Further, although arousal and mood effects have inconsistent effects on attention and performance, it may be that the familiar and unfamiliar excerpts used by Perham and Vizard (2011) induced arousal and mood effects in opposite directions. Although if this is the case that listening to liked music increases positive arousal, their findings do challenge the view that positive mood facilitates task performance (Schellenberg, 2005, see 2.7; see also Threadgold et al., 2019).

In relation to potential disruption produced by irrelevant music with different familiarity, there are at least two factors, aside from arousal/mood and the changing-state of a sound, that also need consideration, particularly in the context of theories concerning attentional capture. These factors are predictability and personal significance. A familiar melody differs from an unfamiliar melody in terms of its predictability. Participants are generally quick to recognize a familiar melody given a few notes (Dalla Bella et al., 2003). While such predictions are also made for unfamiliar melodies in relation to musical structure or syntax (e.g., harmonically expected notes or chords), the predictions are more exact for a familiar melody for which participants possess prior knowledge of the sequence order of pitch/rhythm as well as musical structure. On the attentional capture view, unexpected changes within an auditory stream captures attention. Given that participants should have a more precise neural model for familiar against unfamiliar material, one might argue that unfamiliar melodies should be more distracting than familiar melodies. This expectation, however, has yet to be studied in the context of long-term memory. However, Hughes and Marsh (2019) showed that pre-exposure to streams of changing-state letters, for which a neural model could then be specified, did not reduce the distraction they conferred to serial recall. This suggests the existence of a neural model for familiar material may not aid performance. In contrast, Hughes and Marsh (2019) also found that prior exposure to more complex material (sentences) reduced distraction relative to when no foreknowledge was given, regardless of whether a short-term memory task required seriation. Thus, a reduction in distraction may depend on the specificity of more complex information about the distracter (e.g., its syntax). If so, then one might expect familiar melody to be less distracting than unfamiliar melody, because participants presumably already have a better specified neural model of the syntax or structure of a familiar as opposed to unfamiliar melody. However, if one supposes that nursery rhymes are capable of eliciting person-specific autobiographical memories then they might be perceived as over-familiar or of personal relevance. A personally significant stimulus such as one's own name can also capture attention relative to a yoked control name (e.g., Röer et al., 2013). Therefore, this view, at odds with the predictability view, supposes familiar melody to be more disruptive than unfamiliar melody due to attentional capture. (This idea will be explored further in Chapter V-VI.)

The foregoing has focused on irrelevant sound/music during tasks that require serial rehearsal within the ISE paradigm. However, irrelevant music stimuli have caused noticeable disruption to cognitive task performance not reliant on serial recall such as reading comprehension (Martin, Wogalter, & Forlano, 1998; Perham & Currie, 2014), or mental arithmetic (Perham et al., 2013) and consideration of this literature now follows.

2.8 Music-auditory distraction during non-seriation tasks

Music, as a description for organized sound, is a broad church, and has many facets (such as articulation, genre, mood, or familiarity) that can affect the listener. However, when undertaking cognitive tasks such as reading, or studying, music is frequently processed as a conscious or unconscious associate. An idea frequently voiced in relation to disruption of task performance by background sound is that the processing of the irrelevant stimulus usurps some cognitive capacity that could otherwise be dedicated to focal task processing, resulting in a decline in task performance (e.g., Neath, 2000). This idea has also been espoused in the context of music (Chamorro-Premuzic, Swami, Terrado, & Furnham, 2009; Konecni, 1982) wherein it is assumed that the presence of background music results in a "dual-tasking" situation wherein some limited resource attention is consumed by the music to the detriment of performing any attentionally-demanding focal task regardless of whether seriation is entailed. Capacity-based arguments, however, are circular since resource accounts typically do not explain how a resource or resources are or become(s) limited (Neumann, 1987; see also section 2.5.1).

Whether music impairs or enhances task performance remains the source of perennial debate. There may be a difference between subjective ratings of susceptibility to distraction and acceptability or preference for the sound and its objective impact on task performance. For example, Mowesian and Heyer (1973) in arithmetic/spelling tasks found that although participants preferred a background distraction of rock music compared to folk, classical instrumental, or vocal-operatic, performance of all group scores were similar with the music having no detrimental impact against a quiet condition. In addition, Pool (2002) found pencil and paper memorization task results did not deteriorate if completed during background music as compared to the detrimental effect of a background video. One possibility that background video disrupts performance more is that "cuts" within film may draw visual attention away from the focal task (Pool, Koolstra, & van der Voort, 2003).

The failure of music to disrupt performance for tasks that do not call upon seriation is consistent with the interference-by-process account. In this view, the focal task must call upon seriation processes to be vulnerable to disruption via the changing-state properties of background sound (e.g., Beaman & Jones, 1997). Arguably, the particular arithmetic and spelling tasks used by Mowesian and Heyer (1973) and the pencil-and-paper memorization tasks used by Pool (2002) did not necessitate seriation. Findings that calming, relaxing music, fails to disrupt performance, or can improve performance (Thompson et al., 2001) is consistent with interference-by-process in two ways. First, if the focal task does not involve seriation (as in Thompson et al., 2001) then there is no clash of processes and presumably other effects (e.g., mood/arousal) can operate. Second, calming music with low levels of acoustic complexity might be expected to modulate performance via mechanisms other than an interference-by-process (Jones & Tremblay, 2000).

2.8.1 Reading comprehension/prose/mental arithmetic.

In contrast to other tasks involving spelling (Mowhesian & Heyer, 1973), cognitive demands for reading are considerable (Walczyk, 2000, cited in Thompson et al., 2011). Language and music involve a structure of phonemes/pitch that evolve through time to allow meaningful realisation into phrases/sentences/melody. In other words, in addition to identification of each letter and semantic access there must be a process for understanding and developing the resulting word when reading. Similarly, a sequence of pitch requires a knowledge and understanding of the relationship between individual sounds to identify the resulting melodic phrase (see Chapter I). As both domains are reliant on temporal processing skills competition for processes may result during when a task is undertaken in the presence of background music and more so if the to-be-ignored input is vocal (Thompson et al., 2011). In the context of reading comprehension, findings in relation to performance changes as a function of background music presence are mixed. For example, Hilliard and Tolin (1979) found familiar, as compared to unfamiliar, background music actually enhanced comprehension task performance. Other findings have confirmed genre to be non-influential on task performance, for example, rock and roll on reading comprehension (Tucker & Bushman, 1991). Nevertheless, assertions have been made that "the detrimental effects of background music on reading comprehension likely depends on the characteristics of the music" (Kämpfe, Sedlmeier, & Renkewitz, 2010, p. 702). Consistent with this conclusion, Furnham and Stephenson (2007) identified better performance on reading comprehension, free recall, mental arithmetic, and verbal reasoning, when background music was calm as compared to upbeat. However, an earlier study showed reading comprehension, prose recall, and mental arithmetic, with irrelevant garage music (1990's electronic presumed "upbeat") and office noise auditory distraction, indicated no significant difference between either auditory condition (Furnham & Strbac, 2002).

A further complicating factor in this line of work, however, is that personality-type sometimes mediates the effects of background sound on performance. In their study Furnham and Strbac (2002) showed extroverts performed slightly better than introverts when reading comprehension was accompanied by background music. This may be the result of music's potential arousal-inducing effects (see 2.7). However, such a conclusion is speculative since experiments on auditory distraction do not always take account of participant's arousal/reaction to the varying types of music/genres they are exposed to. Chou (2010) did show how supposed musical mood may impact differentially on reading comprehension in silence and during a "Chill with Mozart" instrumental mix condition, and "Hip Hop Best" condition performed by various male and female vocalists (hip-hop being a form of chanted rhythmic, rhyming speech from 1970's). Although scores were highest in silence, vocal hip hop had the most deleterious effect on reading comprehension.

A pervasive problem in work comparing genres and assumed musical mood, however, relates to acoustic differences between the musical excerpts deployed. For example, in the study of Chou (2010), tempo and rhythmic patterns within the distracter genres are inconsistent as are the mix of timbre. These confounding variables need to be considered, alongside the limited pitch contour (less changing-state) of hip-hop vocals, when deducing effects from their results. It could be that acoustic differences between conditions, rather than assumed musical mood drives different effects. However, another possibility is that the two are perfectly confounded since the very attribute(s) that convey, or induce musical mood, might also be the attribute(s) that produce disruption of task performance regardless of the mood induced.

2.9 Tempo, intensity, dynamic, and motor action in non-serial tasks

In addition to genre, familiarity, preference, or lyrics of background music, other musical elements (notably tempo and dynamic), alone or in combination, may have influential effects on performance depending on focal task requirement. For example, self-selected and experimenter-selected music across two intensities was used to test participant reaction in a driving-game (Cassidy & MacDonald, 2010). Increased engagement was generated by self-selected music (presumed familiar). However, although increased tempo for experimenter-selected music created faster performance, it also resulted in increased inaccuracy. Cassidy and MacDonald (2010) deduced self-selected music may hold personal significance resulting in a more positive focus. It should be considered here that no account was taken as to inclusion of lyrics (Iwanaga & Ito, 2012), or to assess familiarity/unfamiliarity (Parmentier & Andrés, 2010), or personal significance (Baumgartner, 1992; Rathbone, O'Connor, & Moulin, 2017) which may be disruptive. Moreover, personal significance has been shown detrimental to serial recall (e.g., Röer et al., 2013).

Intensity (high/low) and tempo (fast/slow) manipulation of background music (Mozart Sonata in D for two pianos, K448) delivered during reading comprehension showed fast/high instrumental music to be most disruptive on accuracy (Thompson et al., 2011). The element pairing was slow/low, slow/high, fast/low, fast/high, with ranges from 60dB-82dB and 110 bpm-150 bpm. Two groups of participants, one musically trained, read a passage before answering multiple choice questions that involved deep comprehension. Thompson et al. (2011, p. 6) suggested that higher results in slow tempo conditions allowed for "continuous and spontaneous recovery from acoustic interference". (Musicians outperformed non-musicians, supporting intellectual advantage shown by Schellenberg, 2004, 2006.) However, Thompson et al. did not include music with lyrics or varied genres which may have produced different arousal with potential to affect task accuracy. A relationship between tempo of background music and tempo of focal task found to exist whereby tempo can transfer to concurrent motor activity: rhythm of the auditory input appears to influence the task rhythm, e.g., if background music is fast, water is drunk faster (McElrea & Standing, 1992).

2.10 Semantic-auditory distraction

Research in the context of visual-verbal serial recall demonstrated that the acoustic rather than post-categorical, semantic, properties of sound disrupt performance. However, a growing body of research demonstrates that semantic properties of sound assume disruptive power if the focal task draws on semantic processing (Marsh & Jones, 2010). An overview of semantic auditory distraction is justified at this point since in the forthcoming empirical Chapter IV lyrics retrieval and production in the presence of background sound conveying semantic properties, or not, will be addressed. Semantic auditory distraction effects have been cast within the interference-by-process framework (Marsh, Hughes, & Jones, 2008, 2009; Oswald, Tremblay, & Jones, 2000), as they tend to manifest when both irrelevant sound and the focal task exhibit semantic properties. This situation would occur when trying to speak familiar lyrics during different irrelevant spoken- or sung-lyrics (see Chapter IV). However, alternative accounts of semantic auditory distraction are possible (see Marsh et al., 2008) and these will be entertained in the forthcoming sections.

2.10.1 Semantic-auditory distraction and the ISE in serial recall.

For serial recall tasks an irrelevant sound's content is unimportant. For example, the magnitude of impairment to serial recall has been shown to be no different when the sound's meaning could be understood or not (e.g., presented in a foreign language; see LeCompte et al., 1997; Jones et al., 1990; Salamé & Baddeley, 1982). Moreover, the semantic similarity in content between irrelevant sound items and to-be-recalled items does not influence the magnitude of disruption to serial recall (Buchner et al., 1996; Marsh et al., 2008, 2009). These findings are readily accounted for within the interference-by-process framework (Marsh et al., 2008; Marsh & Jones, 2010). Most serial recall experiments have used single digits or consonants. In this setting, the to-be-remembered sequence is stripped of any syntax or semantic information that could support recall. This set of circumstances is what forces participants to use serial rehearsal to impose order cues on a list of items that possesses no intrinsic cues to item order. Since rehearsal is a non-semantic articulatory-driven process for which the vocal-motor system is utilized, the semantic properties of sound are ineffective in producing disruption. Even though there is good evidence that irrelevant sound is automatically processed semantically (Vachon, Marsh, & Labonté, 2020) the lack of semantic processing in the focal serial recall task means that there is no clash of similar semantic processes (Marsh et al., 2008). Regardless of whether it conveys meaningful or meaningless sound, irrelevant items semantically similar, or dissimilar to the to-be-recalled items, disrupts performance due to its acoustic, changing-state properties. The obligatory processing of these acoustic changes, that likely vary only by chance across different types of irrelevant sequences (e.g., meaningful *vs.* meaningless sound and to-be-ignored sequences of items semantically similar *vs*. dissimilar to to-be-recalled items), conflict with the deliberate, non-semantic serial rehearsal processes responsible for maintaining the order of to-be-recalled items.

Consistent with the interference-by-process framework, if the focal task calls on semantic processing, then semantic properties of irrelevant sound assume some disruptive power (Banbury, Macken, Tremblay, & Jones, 2001; Marsh & Jones, 2010). In other words, when semantic analysis and processing is required for the focal task, in the presence of an irrelevant sound that also exhibits semantic properties, semantic auditory distraction effects are generated. These semantic auditory distraction effects are interpreted within the interference-by-process framework as resultant from a semantic-interference-by-process.

2.10.2 Semantic-auditory distraction in non-serial episodic recall.

Although semanticity (meaning) has little influence over the classical ISE in serial recall (e.g., Buchner et al., 1996, see 2.3) some non-serial recall tasks have been detrimentally affected by between-stream semantic similarity (irrelevant sound and task) when concurrent semantic analysis and processing are required (Jones et al., 1990; Oswald et al., 2000). For example, Marsh et al. (2008) presented participants with a list of ten-items wherein each list comprised examples drawn from a single semantic category (e.g., Fruit: “peach”) and requested one group of participants recall the list items in any order (free recall) and another group recall them according to their serial order (serial recall). Irrelevant sound comprising items drawn from the same, as compared with different, semantic category to list items resulted in poorer recall performance only when participants were engaged in free recall. This between-sequence semantic similarity effect also obtained in a later study in which participants were presented with lists of items within which several items from a number of categories were presented in a random order. Recall performance was poorer when participants attempted to recall list items by category (categorization) as compared with by their serial order of presentation (Marsh, Hughes, & Jones, 2009). These two studies also revealed an effect of meaningfulness in that meaningful as compared to meaningless irrelevant sound (e.g., words as compared with non-words, forward speech as compared with reversed speech) produced greater disruption to free recall (Marsh et al., 2008) and categorization (Marsh et al., 2009). These results are consistent with the notion that free recall and categorization processes tap semantic processing and it is this deliberate semantic processing that comes into conflict with the automatic semantic processing of irrelevant sound (Marsh et al., 2008, 2009).

The impact of between-sequence semantic similarity on free recall and categorization, has been considered within a dynamic account of forgetting (Marsh et al., 2008). Specifically, Marsh et al. (2008) followed Anderson and colleagues (Anderson, 2003) in assuming that retrieval from long-term memory functions much like an act of selective attention that is externally focused, with the exception that it is focused on internal representations. When participants probe their memory for previously presented items, they may do so with underspecified retrieval cues. They may have processed the shared semantic relationship between items and use this concept (e.g., “Fruit”) as retrieval cues. However, retrieval is competitive and with the cue being shared between presented (target; “apple”) and non-presented (competitor; “banana”) items, many response candidates vie for retrieval thereby threatening to impede selection of targets. The isolation of a target from a competitor, according to Anderson and colleagues (Anderson, 2003; Anderson & Neely, 1996), requires an inhibitory process that prevents competitors from being retrieved. However, this has a consequence: targets that share some of the same features as competitors are also inhibited, thus impeding their retrieval (Anderson & Spellman, 1995). This view has been extended to disruption via meaningfulness (e.g., due to words as compared with non-words). The notion here is that there is an overhead cost—in terms of performance impairment—from preventing the incorrect set of semantic representations from engaging the semantic-organizational processes that support retrieval (or decoupling the incorrect set of semantic representations from those processes once engaged). This process-oriented view denies that forgetting is produced by an interference between the content of similar representations and supposes instead that it is the by-product of a dynamic and functional process that ordinarily serves to decrease the accessibility of task-irrelevant representations from being retrieved or encroaching upon goal-directed behaviour. Free recall and categorization tasks are important in demonstrating semantic auditory distraction, but arguably of greater relevance to the current thesis—that explores retrieval of song information cued by song-title—is the disruptive effect produced by irrelevant sound on the production of words from long-term, semantic memory (Jones, Marsh, & Hughes, 2012; see also Marsh et al., 2017).

2.10.3 Semantic-auditory distraction in lexical production.

Tasks that involve recall of presented information for lists that are beyond the span of the participant arguably involve both WM and long-term episodic memory. One task that makes minimal demands on WM and episodic processes (such as seriation) is semantic fluency. Semantic fluency tasks involve presenting participants with a category-name and requesting that they generate as many items from that category as possible within a limited time frame (e.g., 60 second, e.g., Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). Information, therefore, is assessed and retrieved from semantic memory in a reasonably direct fashion (Graesser & Mandler, 1978). The semantic fluency task can be performed with success even when undertaken with articulatory suppression (Baddeley, Lewis, & Vallar, 1984) which suggests it is not underpinned by WM processes, although some episodic record (and thus WM capability) might be required to avoid repetition (Rosen & Engle, 1997). Using this multiple-response lexical production task, Jones et al. (2012) reported a disruptive effect of mere meaningfulness and between-sequence similarity from task-irrelevant sound. Specifically, fewer instances of categories were generated in the presence of to-be-ignored words than non-words. Further, participants produced fewer words when the to-be-ignored stream of words comprised examples from a category-associated (e.g., “Vegetables”) against a category non-associated (e.g., “Tools”) with the cued category (e.g., “Fruit”). The authors concluded that disruption to semantic fluency by meaningful sound, or sound comprising items drawn from a category similar to the cued category, was attributable to an interference-by-process. Specifically, the semantic activation of task-irrelevant items conflicts with the spread of activation from category-cue or previously recalled items within the focal task. Similar to previous explanations of disruption to free recall and categorical processes, Jones et al. (2012) argue that semantic auditory distraction to semantic fluency performance also reflects a cost. They propose that irrelevant semantic information as conveyed by sound represents a varying degree of match for the production criteria (generating category-exemplars) and that this information must be inhibited to the extent that it is a context-appropriate match for that same criteria. A to-be-ignored sequence of “Vegetables” as compared with “Tools” produce more disruption to retrieval of “Fruit” because vegetables and fruit share certain semantic features (e.g., they are mostly round, people cook and eat them) and so they meet more of the semantic criteria for production.

The semantic fluency task requires production of multiple instances but this requirement is not essential to produce semantic auditory distraction. Such effects have also been observed in lexical production tasks that require a single response. For example, response times to name colour-words (e.g., “red”) written in black have been shown to be slower if participants concurrently ignore auditory colour word (e.g., “blue”) as compared with a non-colour word (e.g., “short”; Elliott, Cowan, & Valle-Inclan, 1998; see also Cowan & Barron, 1987; and the Stroop effect, Stroop, 1935). Further, a target picture (e.g., “bed”) is more slowly named when it is co-presented with a to-be-ignored visual, or spoken, word as opposed to a non-word (Glaser, 1992). A semantic similarity effect has also been reported whereby a distracter word (e.g., “crab”) related to a picture (e.g., “ant”) produces more interference compared to an unrelated distracter (e.g., “bolt”; see Collina, Tabossi, & De Simone, 2013). One of the leading accounts of this picture-word interference effect is the response-exclusion account (Miozzo & Caramazza, 2003). On this account, written or spoken distracter words gain rapid access to an articulatory output buffer and must be removed before production of a target word can commence thereby producing the delayed production of the picture name. Critical to both the cross-modal Stroop effect and picture word interference is the stimulus onset asynchrony between the distracter word and the to-be-named picture (Abdel, Rahman, & Melinger, 2007; Farrell & Abrams, 2014).

Disruption to task performance that requires recall of semantic category-exemplars from episodic memory and long-term semantic memory by the semantic properties of irrelevant sound (e.g., Jones et al., 2012; Marsh et al., 2008, 2009) are helpful in characterizing semantic interference-by-process. However, semantic auditory distraction effects are also observed in a number of more complex tasks.

Writing and reading tasks.

Tasks that involve written production of texts as well as their comprehension, have been shown to be susceptible to disruption via the semantic properties of irrelevant sound (Martin et al., 1988; Perham & Currie, 2014; Sörqvist, Nöstl, & Halin, 2012). For example, Sörqvist et al. (2012) showed that meaningful irrelevant sound impaired quantitative (e.g., number of characters produced) and qualitative/semantic (e.g., uncorrected typographical errors and generation of propositions) and increased pauses compared to meaningless sound and quiet, while meaningless sound produced no disruption relative to quiet.

The detection of errors in written text—proof-reading—has been found susceptible to disruption via the semantic properties of irrelevant sound. For example, Jones et al. (1990) reported meaningful speech produced more disruption than meaningless speech on detection of non-contextual errors (e.g., misspellings) from visual display units when the display size was large (5 lines at a time) as compared with small (1 line at a time). However, detection of contextual errors (e.g., grammatical errors) was unaffected. Meaningful speech was later found to be no more disruptive than meaningless speech in the context of serial recall thereby supporting the concept that disruption occurs from semantic properties of irrelevant sound only when there is a clash of semantic processing.

In contrast to Jones et al. (1990), Halin, Marsh, Hellman, Hellström, and Sörqvist (2014) reported that detection of contextual errors was impaired by meaningful speech as compared with quiet, but non-contextual error detection was unaffected. Further, they report that this disruption was eliminated when the written material was presented in a difficult to read font or masked by visual noise. This suggests that semantic auditory distraction, like the deviation effect, may be tempered by top-down cognitive control (see 2.5.2). In contrast to disruption produced by meaningful speech, meaningless speech and irrelevant music have been shown to be ineffective in disrupting proof reading (Boyle & Coltheart, 1996). This supports the interference-by-process account in showing that the acoustic properties of sound fail to disrupt semantically-oriented focal tasks.

In the context of reading comprehension Martin et al. (1988) reported decreased performance when irrelevant sound comprised words in a language participants understood, as compared to a language foreign to participants or white noise, regardless of whether those words formed a coherent (prose) or incoherent stream (random words). Although reading comprehension suffered in a meaningless sound condition (e.g., speech foreign to the participant), its disruption was no greater than that produced by steady-state white noise which suggests that the disruption effect produced by meaningless speech was not driven by its changing-state properties nor its phonological content. Further, Martin et al. (1988) found greater reading comprehension disruption produced by irrelevant speech as compared with background music, and from music with lyrics as compared to instrumental music. Martin et al. (1988) concluded that impairment produced to reading comprehension by background sound was due to semantic interference between written and spoken word.

This semantic interference-by-process view (Martin et al., 1988; see also Marsh et al., 2008, 2009), is further supported by Perham and Currie (2014). They report that irrelevant background music with sung-lyrics, irrespective of participant's musical preference (e.g., background music liked or disliked) as compared with no lyrics, impaired accuracy of reading comprehension. However, irrelevant background music without lyrics produced no disruption as compared with quiet. The authors suggested their findings cohered with the interpretation that disruption results from concurrent processing of two sources of semantic information: deliberate processing of visually-presented text for meaning, and automatic semantic processing of meaningful lyrics (e.g., Marsh et al., 2008). Further, they adopted the position that disruption of reading comprehension may arise in part from the process of preventing irrelevant semantic activation carried by the lyrics from encroaching onto the deliberate semantic processes required to integrate word meanings and extract the global meaning of text (Jones et al., 2012; Marsh et al., 2008, 2009).

"Music-reading" comprehension.

Music semantic memory describes familiar/known music (Platel, 2005). Speculatively, irrelevant familiar music could, therefore, impact on tasks requiring semantic processing. Familiar melody has been shown to prompt or cue text (Crowder et al., 1990), and may automatically trigger associated lyrics (Perez et al., 2004; Pring & Walker, 1994; see section 2.6). For example, Peynircioğlu et al. (2008) investigated whether titles, melodies, or lyrics were the best cue for the respective components of song (title, melody or lyrics). They found that recall of titles of instrumental extracts was better when lyrics were presented as cues, as compared with melody. Moreover, lyrics were better cues for melody than were titles. However, despite being the strongest cues for eliciting melody and title, lyrics themselves were not very well recalled, regardless of whether the cue was a title or melody. At first glance this finding seemed to suggest that some qualification is required for the bidirectional automatic activation of connections between lyrics and melodies (Peretz et al., 2004) in that the bidirectionality might not be equally as strong (Peynircioğlu et al., 2008). However, it should be noted that Peynircioğlu et al. (2008) did not use highly familiar songs and thus lyrics retrieval may be poor due to lack of familiarity or knowledge of the song's lyrics. If it is the case, however, that familiar melody activates associated lyrics (Peretz et al., 2004) then it might be surmised that the automatic semantic activation of those lyrics could impair performance of a semantic-oriented task via a semantic interference-by-process between imagined lyrics and deliberate semantic activation of the focal task material. However, there is a lack of empirical evidence for this. Feng and Bidelman (2015) report that familiar music is less distracting than unfamiliar music on a word-pair semantic judgement task that involves comprehension. However, the authors used classical music excerpts that do not have associated lyrics and so the findings cannot be due to semantic interference-by-process between imagined and deliberate processing.

Other studies have sought to investigate whether musical comprehension is affected to a greater degree by irrelevant music compared to irrelevant speech. Brodsky, Henik, Rubinstein, and Zorman (2003) requested participant musicians to identify familiar songs from notated scores (and thus notational audiation, e.g., internally hearing/singing the song required, see 2.6.1). In this setting, irrelevant speech produced much less disruption than irrelevant music and these results conceptually replicate a previous study (see also Martin et al., 1988, Experiment 3). The findings with showing greater disruption by music than speech on notational audiation are important in that they rule out any possibility that speech is more disruptive than non-speech regardless of the focal task, as would be the case if it led to attentional diversion (Cowan, 1995). In contrast, results fit better with the notion that the degree of overlap between processing mechanisms, those required for task completion and those applied to irrelevant sounds, determine the degree of disruption (Jones & Tremblay, 2000; Marsh et al., 2008, 2009; Martin et al., 1988).

Problem solving/mental arithmetic.

A number of studies demonstrate that problem solving is vulnerable to the semantic properties of irrelevant sound. For example, number counting can be impaired when there is semantic similarity between a running total and to-be-ignored items (Buchner, Steffens, Irmen, & Wender, 1998). Further, Graydon and Eysenck (1989) report slower response times when participants are requested to count backwards (in ones from two three-digit numbers) in the presence of task-irrelevant digits than letters, demonstrating that between-stream semantic similarity contributes to disruption. However, the authors report that this between-sequence semantic similarity effect is influenced by task complexity. Finally, Perham, Marsh, Clarkson, Lawrence, and Sörqvist (2016) requested participants perform mental addition in the presence of task-irrelevant ascending order number speech, descending order number speech, and quiet. Although the speech content was identical in both conditions, ascending order number speech produced more disruption of mental addition. The authors report that these findings support the interference-by-process account since ascending digit speech is more similar to focal task processing requirements (e.g., use of subvocalization to count and store immediate outcomes), than descending digit speech. Furthermore, their findings run contra to the prediction of the interference-by-content approach wherein identity, not order, of task-irrelevant digits (ascending *vs.* descending) should drive their disruptive effect on mental arithmetic.

2.11 Structural accounts of semantic-auditory distraction

2.11.1 Interference-by-content.

Interference-by-content approaches assumes that focal task impairment results from structural similarity between to-be-remembered and to-be-ignored items (McGeoch, 1942; Salamé & Baddeley, 1982). On these approaches, disruption to a primary task could result from similarity of *semantic content* between visual information and irrelevant auditory information of auditory distraction with the semantic content of target items. The between-sequence semantic similarity effect could then be explicable as arising as a passive by-product of greater overlap, within a semantic memory store, or semantic psychological space. Such an explanation would assume that to-be-ignored items gain access to the same place as to-be-remembered items within the cognitive system and thereafter their co-occurrence causes either degradation of items, confusion between the items of different origin, or displacement of items from the store or space. The feature model (e.g., Neath, 2000) is a good example of such an approach.

Feature adoption.

Degrading of primary memory traces of to-be-recalled items in the presence of distracters, according to the feature model, may lead to feature adoption and subsequent interference. Feature adoption (see Neath, 2000) is a mechanism whereby features of a to-be-ignored item can alter the features of a to-be-recalled item if the values of those features mismatch (although some temporal correspondence must occur for this to be possible; Neath, 2000). The consequence of feature adoption will be a degraded representation of the to-be-recalled item. Its recall will then be impaired since recall is a process of matching primary memory trace to a correct item in secondary memory. Importantly, such features may represent semantic properties. If features can represent semantic content, then it is possible to envisage how mismatching features would be more likely to be overwritten by semantically-similar than semantically-dissimilar to-be-remembered and to-be-ignored items, because in the latter case there are more features in common. However, it is unlikely that semantic feature adoption can explain semantic auditory distraction. There is good evidence that semantic properties are automatically activated upon written and aural presentation of target and distracter material respectively (Vachon et al., 2020). However, a between-sequence semantic similarity effect only emerges when the task-process requires semantic processing, as in free recall or categorization against serial recall (Marsh et al., 2008, 2009). The feature model would thus have to make a number of additional gyrations to explain why that between-sequence semantic similarity effect is modulated by task-process. Furthermore, it is difficult to envisage how constructs within the feature model could explain disruption to more complex tasks such as reading comprehension (Martin et al., 1988) wherein there is limited semantic similarity between target and irrelevant information (see also Meng, Lan, Yan, Marsh, & Liversedge, 2020).

2.11.2 Attentional accounts.

Several approaches have attempted to explain semantic auditory distraction with reference to attention. For example, some structural approaches suppose performance under distraction becomes limited due to the distracting material drawing upon a limited attentional resource or set of resources. For example, it might be assumed that there is limited capacity for semantic processing within the system and that semantic processing of task-irrelevant material draws on the same resources as to target material (Hygge et al., 2003; Kahneman & Treisman, 1984). However, resource views in general are considered circular, it is logically incorrect to infer from the observation that semantic processing is interfered with by irrelevant information, that this is evidence for a limited resource or processing resources for semantic processing since the latter is an *a priori* theoretical assumption. Similarly, it makes no logical sense to use a notion that there is a limited resource or resources for semantic processing to explain why such processing is disrupted in the presence of meaningful task-irrelevant speech because, again, the notion of a limited resource for semantic processing is an *a priori* assumption (for further elaboration, see Allport, 1993).

Another attentional account that is prey to many of the same critiques as general resource based accounts (above) is the attentional capture account (Cowan, 1995). Although attentional capture is most often characterized as resulting from processing of acoustic novelty (Cowan, 1995), one possibility is that primed semantic features produce attentional capture. Stimuli that have a high resting activation level within memory can have the potential to be triggered beyond an activation threshold that would capture attention such as one’s own name (Moray, 1959). In the case of between-sequence semantic similarity, semantic properties of to-be-remembered items may prime semantic features within the sound, thereby producing a semantic attentional capture. However, such an explanation might be a stretch since semantic processing of written word (that might prime semantic features of task-irrelevant items) often occur full-blown regardless of the orienting task (Neely & Kahan, 2001). This does not chime with the finding that a between-sequence semantic similarity effect emerges when the orienting task entails free recall (Marsh et al., 2008) and categorization (Marsh et al., 2009), that encourages semantic processing, but not when the task requires a non-semantic, seriation process (Marsh et al., 2008, 2009). Generally then, attention based accounts fail to explain why the appearance of semantic auditory distraction effects, including those of between-sequence semantic similarity, is dictated by the processing specificity of the prevailing mental activity.

2.12 Interim summary

In sum, it would appear in serial recall that the greater acoustic variability of irrelevant sound the greater disruption to primary task, and in tasks that appear susceptible to semantic-auditory distraction they themselves require semantic processing. This lends credence to the interference-by-process view whereby a conflict is generated between two semantic processes (Jones et al., 1990): familiar, as compared to novel/unfamiliar, auditory distracters able to activate serial-order representations to a greater extent. However, in order to produce/perform song/speech, the vocal-motor production mechanism, required to plan/check output may be vulnerable to pre-attentive processing of serial order within distracter melody/lyrics (Hughes & Marsh, 2017). For example, by perceptually-organizing sound into streams (see Hughes & Jones, 2005). Alternatively, familiar auditory distracters may be more likely to produce *attentional diversion.* These issues are to be addressed in the following empirical chapters that will now be introduced.

Prologue to empirical chapters.

The four empirical chapters within this present thesis constitute a systematic study that explores and evaluates retrieval processes of known song from long-term memory *and* production of the two foremost components, melody and lyrics, within the context of different irrelevant distracter familiar and unfamiliar melody/lyrics manipulations. In addition, the same irrelevant sound conditions are delivered during short-term memory seriation and non-seriation tasks. Therefore, this thesis will look both to and beyond the serial recall paradigm and explore retrieval requiring melody performance and production of spoken-lyrics in tasks that tap music semantic memory and lexical retrieval. In doing this, the studies assess the veracity of competing accounts of auditory distraction: interference-by-content and interference-by-process. These studies will allow investigation of any disruption of the vocal-motor system by the presence of task-irrelevant sequences (perceptual-gestural view, e.g., Hughes & Jones, 2005) and whether the interference-by-process account can be extended from short-term serial recall to long-term memory tasks requiring song-element production. Specifically, whether the construct of interference-by-process can explain auditory distraction in the context of tasks necessitating retrieval of familiar melody/lyrics or tasks requiring short-term retention of verbal material. Production requirements for semantic focal tasks (musical and verbal), throughout irrelevant sounds that also demonstrate acoustic variability and semantic properties, may support this view, and show this account is not just reliant on serial recall. A general overview of the issues to be addressed will now be presented.

Empirical Study I (Chapter III).

Study I will explore and compare familiar melody retrieval and production, via humming, in relation to the presence of various acoustic properties of auditory distraction, namely different melody, sung-lyrics, and spoken-lyrics, familiar and unfamiliar. Here, the key aim is to address if familiar melody will produce more disruption to target melody production than unfamiliar melody, as it is hypothesised their well-formulated motor-plans will generate increased competition with target melody production. In addition, if humming production entails activation of the music lexicon (Peretz & Coltheart, 2003) this would indicate semantic retrieval. It will be expected, therefore, if melody and lyrics within song are stored separately to verbal information, that sung-lyrics will disrupt melody retrieval to a greater extent than spoken-lyrics.

Empirical Study II (Chapter IV).

Study II focuses on lyrics retrieval. Prior evidence regarding whether lyrics and melody within song are processed differently (Sammler et al., 2010), stored together, (e.g., Hébert & Peretz, 1997), or separately (e.g., Berz, 1995; Racette & Peretz, 2007), and whether retention of melody in autobiographical memory can support retrieval of lyrics (Rainey & Larson, 2002; Weiss, Trehub, & Schellenberg, 2012) is still debated. Based on a modular model of music processing (Peretz & Coltheart, 2003) lyrics retrieval from long-term memory, and production via speaking, requires semantic processing that may conflict with concurrent semantic processing of irrelevant lyrics/melody. Study II will enable a direct comparison with Study I regarding the flow of communication between processing components according to the modular model of music processing (Peretz & Coltheart, 2003). A primary hypothesis is that if results from a lyrics retrieval task, during the same auditory distracter conditions as used for melody retrieval (Study I, Chapter III), show the same pattern of disruption this will support the notion that, within song, melody and lyrics appear to be integrated (e.g., Schön et al., 2010). Alternatively, a different pattern of results from those of Study I will point to some degree of separation (e.g., Bonnel et al., 2001). A second hypothesis is that spoken- or sung-lyrics will impair spoken-lyrics performance to a greater degree than melody alone, the semantic rather than the phonological properties of irrelevant sound more influential on task accuracy (Beaman, 2004; Marsh et al., 2009). Disruption to semantic retrieval, like serial recall, will therefore, be investigated by examination of the interactions of motor-systems and dynamic processes.

Empirical Study III (Chapter V).

The broad aim of Study III is to investigate the interference-by-process proposition by comparing results from a serial short-term memory task with the same auditory distracter conditions as Study I-II. On the interference-by-process view (Jones & Tremblay, 2000), it is hypothesised that the changing acoustic properties of irrelevant sound should disrupt serial recall accuracy, although speech, being more acoustically complex than melody, should produce greater disruption. In addition, there should be no difference in recall accuracy with familiar as compared to unfamiliar distracters. However, according to the specific attentional capture view of the duplex account (Hughes, 2014), attentional capture might be task-independent, and should manifest in a serial recall task wherein retrieval of melody/lyrics is not required. Therefore, the interference-by-process view is pitted against an attentional capture view.

Empirical Study IV (Chapter VI).

The main aim of Study IV is to explore whether any additional disruption due to music familiarity is due to attentional capture or enhanced seriation (due to a well formulated neural model of the sequence). This is explored within the context of a missing-item task because (arguably) it does not involve seriation but is susceptible to attentional capture effects. Distraction effects (e.g., familiarity) reported in the context of song production will be qualitatively distinct to that found in serial recall tasks. Confirmation of these hypotheses would yield a pattern of data consistent with the notion that the same generic principles underpinning distraction in serial recall can also apply to retrieval and production of melody/lyrics from semantic memory.

CHAPTER III

EMPIRICAL STUDY I: DISRUPTION OF MELODY RETRIEVAL BY IRRELEVANT SOUND: INTERFERENCE BY PROCESS?

Abstract

The susceptibility of selective attention across modalities can be illuminated through investigating distraction produced by background sound on performance of visually-based cognitive tasks. Study I centred on auditory distraction during tests of memory retrieval for familiar songs. Participants were required to hum the melody cued by song-title while exposed to auditory distracters comprising familiar or unfamiliar melody, with or without sung-lyrics, or spoken-lyrics. It was hypothesized that familiar melody would produce more disruption than unfamiliar melody by activating competing melodies within long-term memory and yielding well-formulated competing motor-plans. Results showed unequivocally that both familiar and unfamiliar distracters impaired retrieval accuracy. Furthermore, spoken-lyrics and melody, regardless of whether familiar or unfamiliar, impaired performance to a lesser extent than familiar lyrics and melody combined. This pattern of results reflects a legacy cost associated with the decoupling of a familiar irrelevant melody from the vocal-motor system enabling a target melody to be produced. Moreover, it is argued that this pattern of results speaks to a current debate concerning whether lyrics and melody are integrated or separable. Familiar melody without lyrics was not significantly more disruptive than unfamiliar melody without lyrics suggesting that lyrics were not automatically activated by the familiar melody *per se*. However, the fact that familiar melody and lyrics combined (sung-lyrics) was the most potent distracter suggests that humming of a target melody may invoke retrieval of associated lyrics rather than some peripheral activation of the motor-system. These results favour a process-oriented, approach to attentional selectivity during retrieval of melody.

3.1 Introduction

Hearing is a unique sense. Irrespective of how we are otherwise engaged, an irrelevant auditory stimulus can impact on our goal-directed behaviour. Selective attention must be permeable so that a necessary response can be made—for example, to alert danger—irrespective of current behaviour. However, it is this quality that leaves an organism open to the necessary consequence of distraction (Hughes & Jones, 2003). It has been empirically demonstrated numerous times that task-irrelevant sound that participants are instructed to deliberately ignore has the capacity to infiltrate selective attention to the detriment of cognitive task performance (Beaman & Jones, 1997; Bell, Buchner, & Mund, 2008; Cassidy & MacDonald, 2007, 2009; Colle & Welsh, 1976; Furnham & Strbac, 2002; Hughes et al., 2007; Jones et al., 1992; Jones & Tremblay, 2000; Martin et al., 1988; Röer et al., 2015; Salamé & Baddeley, 1982, 1986, 1989). The empirical platform upon which the majority of work on auditory distraction has been undertaken uses a verbal-visual serial short-term task (see Chapter V; Conrad, 1964). Within that setting, the presence of to-be-ignored background sound during presentation and the onset of recall, produces appreciable disruption to memory for the correct sequence of verbal-visual items (e.g., Colle & Welsh, 1976; Hughes et al., 2007; Jones & Macken, 1993; Marsh et al., 2009).

Within the domain of visual-verbal short-term memory a debate has ensued between competing accounts of this "irrelevant sound effect". Generally, there is a tripartite classification of such accounts. According to the interference-by-process account, disruption produced by irrelevant sound on visual-verbal serial recall is attributable to a clash between two contemporaneous serial-ordering mechanisms: one deliberate and applied to the focal task—serial rehearsal—and one automatic and applied pre-attentively to the irrelevant sound (Jones et al., 1996; Tremblay & Jones, 1998). The interference-by-content account has a number of manifestations (e.g., Neath, 2000; Oberauer & Lange, 2008; Salamé & Baddeley, 1982) and generally assumes that co-occupation of items from visual and auditory origin within the same store or memory space results in poorer recall of visual, to-be-remembered items. In contrast, attentional diversion accounts come in two guises. On the one hand, according to the feature model (which is also an interference-by-content account; see Neath, 2000), an assumption is made that processing attended and unattended, irrelevant, stimuli simultaneously reduce attentional resources available for processing a focal task. On the other hand, the attentional diversion account (Cowan, 1995, 2001) proposes that changes within sound capture attention away from the focal task to the detriment of task performance (these attentional accounts will be reviewed within the General Discussion section). This current chapter seeks to determine whether the interference-by-process account can be extended to a hitherto unstudied research domain involving retrieval of a known melody in the presence of potentially distracting background sound. The interference-by-content approach is considered as a theoretical counterpoint to the interference-by-process account.

3.1.1 Passive *versus* dynamic view of the interference-forgetting relationship.

Interference within memory research refers to difficulty in remembering or forgetting of stimuli or events. Such interference is thought to arise to the rate that a previously, simultaneously, or subsequently encountered stimulus or event (a competitor) is similar to the to-be-recalled event (the target; Anderson, 2003; Anderson & Neely, 1996; Baddeley, 1986; McGeoch, 1942; Mensink & Raaijmakers, 1988; Salamé & Baddeley, 1982). However, set against this view is one that assumes forgetting is functional. Here, difficulty in remembering, or forgetting, stimuli or events is due to the action of dynamic and adaptive processes of selective attention that resolve interference that arises from stimuli or events that compete with targets during the act of retrieval (Anderson, 2003; Anderson & Bjork, 1994). One example of selective attention process that resolves interference is inhibition (e.g., Anderson, 2003; Tipper, 2001). Inhibition has been considered to serve multiple functions within the system from sensory level to cognitive level. Generally, inhibition encompasses suppression of stimulus components and formation of object representations that underpin selection-of-action. The notion is that initial selection of the target from distracter entails inhibition of representation of the distracter from mechanisms of response output. This can result in a delay in reaction time (e.g., Tipper & Cranston, 1985), or impairment in mnemonic recall (Hughes & Jones, 2005) when the distracter then becomes a target.

A clash between structural and dynamic views of the relationship between interference and forgetting can be observed in microcosm in the context of the irrelevant sound paradigm (e.g., Colle & Welsh, 1976; Jones et al., 1992; LeCompte, 1994; Macken et al., 1999; Salamé & Baddeley, 1982). Within this context, most of the evidence undermines the view that the classic irrelevant sound effect is produced as a passive consequence of similarity between to-be-remembered and to-be-ignored items (Neath, 2000; Salamé & Baddeley, 1982; see Jones & Tremblay, 2000). An early view was that overlap in phonemes between to-be-remembered and to-be-ignored items was an important determinant of the magnitude of the irrelevant sound effect (Salamé &

Baddeley, 1982). This phonological store account proposes that there is a clash between auditory-verbal information that automatically and directly enters a storage module responsible for temporary maintenance of verbal information—the phonological store—and visual-verbal input that gains indirect access to the same store through a grapheme-phoneme conversion process. Herein the degree of disruption produced was assumed to be a function of the magnitude of overlap in phonological features between to-be-remembered and irrelevant items. However, this view has been undermined by a raft of studies demonstrating an absence of a between-stream phonological similarity effect in serial recall (Bell, Mund, & Buchner, 2011; Hughes & Jones, 2005; Jones & Macken, 1995; Larsen et al., 2000; LeCompte & Shaibe, 1997; Marsh et al., 2008; but see Eagan & Chein, 2012). Moreover, irrelevant tones produce qualitatively similar disruption to irrelevant speech (Divin, Coyle, & James, 2001; Elliott, 2002; Jones & Macken, 1993; Sörqvist, 2010a) suggesting that the presence of phonemes *per se* within the irrelevant stream is neither necessary nor sufficient for an irrelevant sound effect to be observed.

The concept of interference-by-content is also embedded in the feature model (Neath, 2000; Neath & Nairne, 1995). According to this model there are two loci of auditory distraction. First, the model holds that similar features within to-be-ignored auditory stimuli—when it conveys speech—overwrite corresponding features within to-be-remembered items whilst those items are being rehearsed. Second, it is assumed that contemporaneous processing of two stimuli sets, only one to-be-attended, reduces the overall attentional resource available for the mnemonic task. However, there are manifold reasons for disaffection with the feature model. For example, the feature model accounts for additional disruption of a sequence of acoustically changing sounds (e.g., G, K, D, L; or a sequence of tones changing in frequency from one to the next) relative to steady-state sounds (e.g., G, G, G, G; or the same tone repeated; the "changing-state effect"; Jones, et al., 1992) entirely through manipulation of the attentional construct which has been considered rather ad hoc.

Structural accounts that emphasize interference-by-content as the most potent causal factor of disruption produced by irrelevant sound enjoy little support from empirical evidence the majority of which contradicts the approaches. Having concluded that the weight of evidence is against the interference-by-content view, focus now falls on a dynamic interference-by-process view of auditory distraction. Interest in this present chapter centres on the possibility that the same principles underpinning distraction in serial recall can also apply to retrieval of melody from semantic memory.

It will be argued that disruption from irrelevant sound in melody retrieval, like serial recall, may be understood through interactions of general-purpose perceptual and motor-systems and dynamic processes whose function is to facilitate goal-directed selective attention. A further goal of the current chapter is to provide insight into representation of music and lyrics within the cognitive system.

3.1.2 The perceptual-motor account and interference-by-process.

An emerging view within cognitive psychology is that processes and systems that are not specifically dedicated to memory, are exploited for mnemonic performance. For example, in the context of verbal short-term memory, performance is conceived by some, as tied to comprehension and production skills of the language system (Acheson & MacDonald, 2009a, b; Buchsbaum & D'Esposito, 2008; Hughes, Marsh, & Jones, 2009; Hughes et al., 2011; MacDonald & Christiansen, 2002; Melby-Lervåg & Hulme, 2010). Similarly, the perceptual-motor account proposes that short-term memory phenomena are explicable by appeal to general-purpose deliberate processes involved in producing coherent sequential actions (such as vocal actions; e.g., Rosenbaum, 2009) and involuntary, pre-attentive processes such as perceptual organization of sequential sounds when auditory stimuli are involved (Bregman, 1990; Sussman, Bregman, & Lee, 2014). The motor component of the perceptual-motor account suggests that a motor-action emulator operates as a sequence output-planning (or rehearsal) device that can be run in emulation mode without production of an overt action. This view gels with work on motor-control that supposes that a "forward model" of an action (comprising instructions to effectors and sensory consequences of those commands) is generated to enable fluent execution of a motor-action and its comparison with the intended action (e.g., Grush, 2004; Shubotz, 2007). While several key serial recall phenomena have been recast in terms of perceptual organization and motor-planning (Hughes, Chamberland, Tremblay, & Jones, 2016; Hughes, Marsh, & Jones, 2009, 2011; Jones et al., 2004, 2006, 2007; Macken, Taylor, & Jones, 2014; Macken et al., 2016; Maidment & Macken, 2012; Maidment, Macken, & Jones, 2013; Woodward et al., 2008), focus here is on its explanation of the irrelevant sound effect (Hughes & Marsh, 2017; see Chapter II, section 2.2 for further discussion).

The perceptual-motor view characterizes verbal serial recall as an under-specification or action-parameters problem (Neumann, 1987, 1996). Since no long-term unitized representation of a random sequence of items exists, participants co-opt the skill of inner-speaking to bind together items that otherwise have no pre-existing sequential ordered relation—e.g., syntax or grammar—to one another. Thus, the function of inner speech is not to refresh items that are decaying within a mnemonic store/space (e.g., Baddeley, 2007) but to place to-be-remembered items on the common carrier of speech such that they can become embodied within a temporarily extended, single, motor-plan. Speech (either overt or emulated via a forward model) serves this function of binding unrelated items together since it is necessary sequentially (due to emulation of the movements, and possession of only one vocal tract). Continuous, prosodic, co-articulatory and paralinguistic characteristics of natural speech (e.g., Sternberg, Wright, Knoll, & Monsell, 1980) can further constrain the serial order of items by instilling cues within the motor-plan (Hughes et al., 2009; Macken et al., 2014; Maybery, Parmentier, & Jones, 2002; Neisser, 1967; Woodward et al., 2008).

On the perceptual-motor view, susceptibility of the motor-sequence plan underpinning speech, to distraction by irrelevant sound, stems from under-specification of its action parameters. More specifically, speech, as a skill, is an abstract entity: The guidance of particular behaviours is specified by generic sets of action parameters (that govern movements of the vocal tract to produce coherent speech) but that general set of action parameters is initially unpopulated in terms of specific to-be-remembered content (e.g., phonemes, words, phrases etc., Hommel, 2010; Neumann, 1987, 1996). Non-specificity of the action parameters of speech—or its need to be populated—leaves it vulnerable to interference by any prospective irrelevant input that, although not specifying task-appropriate action-parameters, can populate the motor-plan. Therefore, in this approach, limitations in performance on verbal serial recall tasks in the presence of irrelevant sound do not reflect evidence that mnemonic capacity is limited due to the existence of items that are prone to decay, interference, or displacement with a static short-term memory store (e.g., phonological store) or space (see Hughes et al., 2011; Reisberg, Rappaport, & O'Shaughnessy, 1984).

According to the perceptual-motor view, serial recall performance is impaired by irrelevant changing-state sequences more than irrelevant steady-state sequences because the integrity of the motor-planning process is threatened by the obligatory process involved in perceptually-organizing sound into streams (e.g., Hughes & Jones, 2005; Jones & Macken, 1993). On this process, successive sounds that are acoustically similar to one another (e.g., in terms of frequency or timbre) and therefore share a common ground, are assigned to the same stream and thus computed as belonging to the same environment event (Bregman, 1990; Warren, 1999). Providing they are similar enough to one another, when sound changes from one to the next, they are assigned to the same stream and their serial order is easily discerned. However, successive sounds that differ considerably from one to the next become partitioned into distinct streams and, as a consequence, the veridical order of those successive sounds are difficult to perceive (e.g., alternating speakers; Hughes et al., 2009, 2011; or alternating tones of different frequency; Bregman & Campbell, 1971). The strong serial-ordered sequence formed from assigning successive sounds to the same stream then competes as a candidate for populating the motor-plan and thereby disrupts serial recall (i.e., the changing-state effect, e.g., Jones & Macken, 1993).

The perceptual-motor view thus readily explains why sequences of pure tones (Divin et al., 2001; Elliott, 2002; Jones & Macken, 1993, Sörqvist, 2010), non-vocal music (Klatte et al.,1995; Perham & Vizard, 2011; Salamé & Baddeley, 1989; Schlittmeier et al., 2008), pitch-glides (Jones et al., 1993; Klatte et al., 1995), and noise-bursts (Tremblay et al., 2001) can impair verbal serial recall, since distracters need not be verbal to be perceptually streamed. In addition, this view easily explains why streaming modulates the irrelevant sound effect when the same content is presented, but across different ears (see Chapter II, section 2.5.1; Jones et al., 1999).

3.1.3 Interference-by-semantic process.

While the interference-by-process account has enjoyed much success in the domain of serial short-term memory, it has also been extended to tasks involving semantic processing with interest here hinging on recall of semantic representations from long-term episodic (Beaman, 2004; Beaman et al., 2015; Marsh et al., 2008, 2009, 2012, 2014, 2015a, b) or semantic memory (Jones et al., 2012; Marsh et al., 2017; see Chapter II, section 2.13). In the context of long-term episodic memory, a typical finding is that ability to recall target words from a familiar target category (e.g., “Fruit”) is impaired more by presenting words from the same category as irrelevant sound items compared to when those auditory distracters are taken from a different semantic category (e.g., “Occupations”; Beaman, 2004; Marsh et al., 2008, 2012, 2015a, b; Neely & LeCompte, 1999). Although a number of potential explanations for this between-sequence semantic similarity effect exists (Hanczakowski, Beaman, & Jones, 2017), one prominent explanation supposes that disruption reflects an overhead cost that is incurred when inhibitory processes are recruited to prevent processing of the distracters (Marsh et al., 2008, 2012). Impetus for this inhibitory account (Marsh et al., 2008, 2012) came in part from the work of Anderson and colleagues (e.g., Anderson & Neely, 1996).

Anderson and Neely (1996) have likened retrieval of semantic information from episodic (long-term) memory to an internally focused selective attention act. Memory search is initiated with retrieval cues that are under specified (e.g., “Tools”). As a consequence of this under-specification, a number of response candidates (e.g., “hammer”, “screwdriver”) may become active thereby competing with, and impeding selection of, the target (e.g., “drill”). Those non-target items are called competitors to reflect their action on retrieval of the target in memory search. Isolation of the target response from its competitors is argued to entail the executive process of competitor-inhibition. While this can aid target recall it also entails a consequence since other items, sometimes other target responses, also get inhibited to the extent that they share the same features from the current target which results in their future retrieval impairment (Anderson, Green, & McCulloch, 2000; Anderson & Spellman, 1995).

In the context of episodic recall of semantic representations (Marsh et al., 2008, 2012), the notion is that inhibition is directed to semantically-related auditory distracters to avoid their further processing and therefore minimizing competition they produce for successful encoding and retrieval of target items. Marsh et al. (2012) demonstrated that when related to-be-ignored auditory items become visually-presented to-be-remembered items on the next memory list, then memory performance is impaired suggesting that those items were previously inhibited as a consequence of them being competitors of previously presented to-be-remembered items (Marsh et al., 2012, 2015; for a different interpretation of these findings, see Hanczakowski et al., 2017).

Focus in the current chapter is on whether the interference-by-process mechanism of the perceptual motor account, coupled with dynamic selective attention processes (e.g., inhibition) can be applied to a novel setting of retrieval of song from long-term memory during task-irrelevant distracter material. It will be argued that principles of interference-by-process and the dynamic cognitive control over auditory distraction extend to settings wherein retrieval of well-known song is involved.

3.1.4 Embodiment of music.

One view gaining popularity is that perception of music involves action simulation (e.g., Godøy & Leman, 2009; Leman, 2007; Leman & Maes, 2014, 2015). This approach, cohering with an action-oriented focus emerging in cognitive science (e.g., Hughes et al, 2009; Prinz, Beisert, & Herwig, 2013), proposes that the human motor system (including gesture and body movements) feature centrally in perception of music (Leman, 2007; Sievers, Polansky, Casey, & Wheatley, 2013). As Leman (2007) proposes:

“Embodiment assumes the existence of mirroring processes that facilitate the encoding of expressive gesture into sounds, and the decoding of sounds into expressive gestures…mirroring implies that the underlying mechanisms for music meaning formation are rooted in sensorimotor principles" (Leman, 2007, p. 237).

One interpretation of this view is that music perception activates motor systems that permit reproduction of melodies such as the perceptual-motor-vocal system and other systems attuned in some individuals with musical expertise, such as finger-gesture.

"When perceiving music, sensorimotor mechanisms are activated that in turn drive the perception of music" (Leman & Maes, 2015, p. 242).

Furthermore, the proposal that forward internal models operating to enable sensory outcome predictions of planned or imagined actions—sensory-auditory-visual—can influence online processing of external sensory information (e.g., music) within the environment (Halász & Cunnington, 2012; Maes & Leman, 2013; Schutz-Bosbach & Prinz, 2007; Witt, 2011) suggests that an intricate relationship exists between motor-planning and musical perception, similar to the relationship between perceptual-organization and motor-planning (e.g., Hughes et al., 2016). Although self-perception during music performance has had limited attention, musical fluency relies on the pairing of action with auditory feedback "recurrent auditory information" (Pfordresher, 2006, p. 195) recent empirical evidence has indicated perception and action for music performance share a common representation (cf. Hommel, Müsseler, Aschersleben, & Prinz, 2001) rather than learned association (e.g., between pitch and finger movement on a keyboard, Pfordresher, 2006).

Additional evidence in support for embodiment of melody/music has been derived from studies of auditory imagery. Several studies have shown that listening to music and imagining it—when the music stops—produces the same neural activity (e.g., Halpern & Zatorre, 1999; King, 2006) depending on the song's familiarity (e.g., Kraemer et al., 2005). Moreover, subvocalization has been implicated in rehearsal of auditory images (Reisberg et al., 1989). Preventing subvocalization (inner voice) has consequence for representation of auditory material in the absence of external stimulation (inner ear). For example, Smith et al. (1995) found that requesting participants to continuously produce an irrelevant utterance (tah) while adjudicating whether there was a rise or fall in pitch from the 2^{nd} to 3^{rd} note within a to-be-imagined familiar melody, impaired performance relative to a condition that did not involve articulatory suppression. Presence of irrelevant speech had the same effect suggesting auditory imagery judgements require sub-vocal rehearsal. Similarly, Baddeley and Andrade (2000) found that a secondary task entailing counting from 1-10, as compared with tapping a pattern on a keyboard, reduced vividness of auditory images reported by their participants. Finally, the number of involuntary experienced unwanted musical thoughts ("earworms"), are reduced by chewing gum that impairs use of the sub-vocal motor process (Beaman et al., 2015). This demonstrates a key role for articulatory motor programming in the generation and "hearing" of involuntary imagery.

From the foregoing it is suggested that both production of a target song and the mere perception of a to-be-ignored distracter (e.g., song) involve some degree of activation within the motor system. Moreover, since known melodies are well-represented in long-term memory, such motor activation would be more extensive and specific for familiar musical sequences as compared to unfamiliar musical sequences.

3.1.5 Dynamic retrieval of song from the music lexicon.

The notion that retrieval of one song should be impaired by the concurrent presence of another is intuitive, and acknowledged within the literature on inhibition: *"…it is surely more difficult to bring to mind a particular tune if another is being heard…"* (Watkins & Allender, 1897, p. 565, emphasis in italics added).

However, no studies appear to have addressed whether song retrieval is impaired by the presence of a to-be-ignored song and whether familiarity with that to-be-ignored song exacerbates any retrieval impairment. Selection of a target song for production entails its isolation from others within the musical lexicon (Peretz et al., 2009; Peretz & Zatorre, 2005). Access to the musical lexicon, however, is automatic (Peretz et al., 2009), meaning that upon detection, task-irrelevant songs can activate competing melodies that can impair selection of the target song for production. Thus, activation of a competitor song when attempting to retrieve a target song constitutes a situation analogous to internally-focused selective attention problem as outlined by Anderson and Neely (1996) wherein the executive process of competitor-inhibition is required to isolate the target response, in this case song, from the competitor song. Note that on this approach, competitors must have a representation within long-term memory to impair retrieval of a target. Thus, well-known, or overlearned, familiar distracter songs should produce more impairment to retrieval of target songs than unfamiliar distracter songs.

Furthermore, due to purported strong links between perception and action within music cognition (Godøy & Leman, 2009; Leman, 2007; Leman & Maes, 2014, 2015), presentation of an irrelevant song could entail that song threatening to (or actually assuming) control (of) motor-programming responsible for target song production. Retrieval of a target song requires representation of a combination of pitch, rhythm, and structure within a temporarily extended, sequential object. Production of song, therefore, requires motor skills necessary to produce the audible item, and perception and feedback relating to the item's outcome (Tsang et al., 2011). According to the vocal sensorimotor-loop singing model (Dalla Bella et al., 2011) the production process begins with retrieval from memory of various elements of the music to be sung, such as pitch and duration. These are then directed to the motor control to facilitate sound output via singing (a process of auditory feedback is also necessary to allow for comparison and correction to output). The vocal-motor production mechanism, therefore, plays a central role in planning for song output as it does for speech (Hughes & Marsh, 2017). As in the case of the irrelevant sound effect then, the motor-planning system—for song production—may be vulnerable to pre-attentive processing of serial order within distracter melodies. Moreover, song familiarity may be an important ingredient of the magnitude of disruption. A familiar melody has a well-known temporarily organized sequential pattern that might be expected to threaten to control action (motor-programming) more than an unfamiliar melody.

3.1.6 Representation of melody and lyrics may determine distraction.

Returning to the phenomenological observation that attempting to produce a tune is more difficult when at the same time ignoring another, we might question what boundary conditions must be met to determine this effect. Presumably how well known the target tune for retrieval could factor in its success of retrieval. But, properties of the distracter tune, for example, its familiarity, or if the melody or its lyrics are more dominant in determining disruption, is questionable. Earlier it was noted familiar melodies, that have a representation in long-term memory, are competitors. However, the magnitude of disruption of retrieval produced by familiar song may depend on whether memories for song entail an integration or separation of lyrics and melody and whether lyrics and melody are presented together within the distracter melody.

Current debate surrounds whether lyrics and melody in song are represented separately or together within memory. Studies of melody recognition demonstrate that song melodies are recognized better when heard with their original words as compared to the text of another equally familiar song (Bonnel et al., 2001). Similarly, words of songs are better recognized when they are heard with the original melody than when they are heard with a different but equally familiar melody (Crowder et al., 1990; Samson & Zattore, 1991). Moreover, an accumulation of data from neuropsychological studies suggests that lyrics and melody are integrated (e.g., Hébert & Peretz, 1997; Samson & Zattore, 1991; Serafine et al., 1984; Serafine et al., 1986). For example, Serafine et al. (1984) demonstrated better recognition for an old melody paired to its existing text as compared with an old melody only, a new melody, an old melody paired with new lyrics, and a new melody paired with old lyrics. Here then, recognition of one element was facilitated by the simultaneous presence of another thereby suggesting the two were integrated in memory. In this context music may act as a retrieval cue to evoke to-be-remembered information (Rainey & Larson, 2002). A melody may be retained in long-term memory while spoken words may disappear within minutes, (Weiss, et al., 2012). In addition, if lyrics are associated with a melody one has been shown to have capacity to trigger the other (Crowder et al, 1990; Koelsch et al., 2002). Peretz et al. (2004) suggested the rhythmic qualities of melody and associated lyrics complement each other and contribute equally to memory for song.

3.1.7 Melody retrieval performance via humming.

Retrieval performance of complete target melodies (nursery rhymes) required for Study I of this empirical series is via humming which is inherently difficult, and a cognitive phenomenon that has received scant attention. In fMRI studies humming has revealed less right hemispheric activity than singing (Özdemir, Norton, & Schlaug, 2006), although vocal performance, irrespective of medium, activates a bihemispheric network. Humming, or "intoned speech", is a vocal device that emits a continuous droning sound which is focused on pitch, requiring no enunciation or phrasing. The rationale for using humming in this study, is that it requires pitch and temporal organization and interaction with the musical lexicon—much like singing lyrics—but does not necessitate lexical production: arguably it does not require access to the phonological lexicon thereby proceeding directly to vocal plan formation (Peretz & Coltheart, 2003). In addition, a familiar melody, normally associated with lyrics, can be recalled, and hummed in the absence of lyrics retrieval. Humming, that arguably tolerates a separation of melody and lyrics in performance, in contrast to overt sung-lyrics, may therefore be considered an advantageous medium of retrieval as it lessens linguistic demands (Dalla Bella & Berkowski, 2009).

Study I addressed the question: Are retrieval and production of familiar melodies disrupted more by familiar or unfamiliar aspects of task-irrelevant song? Using distraction as the key theoretical tool, the focal task requirement was to retrieve and produce a melody, by humming, in response to a visual song title e.g., "Oranges and Lemons", whilst ignoring auditory distraction of another melody, a cappella song, or speech (spoken-lyrics) e.g., "Ten green bottles". Here, the core idea is that vocal production of target memories can be impaired by non-target memories that are activated by the presence of task-irrelevant sound (Marsh et al., 2008, 2009). From the standpoint that verbal retrieval is vulnerable to distraction from activation of similar information stored within semantic memory (Jones et al., 2012) it is hypothesized that production of a familiar melody will be more impaired by the presence of a task-irrelevant familiar melody—that has a representation within long-term memory and can therefore compete with the target melody for retrieval—than an unfamiliar melody combining a novel pitch contour with a familiar rhythmic pattern. However, it is also hypothesized that unfamiliar melody, acting as an irrelevant auditory distraction, will produce some disruption as compared with quiet (Jones & Macken, 1993; Salamé & Baddeley, 1989). If humming does not require access to the phonological lexicon (cf. Peretz & Coltheart, 2003), then it is expected that an irrelevant familiar melody with sung lyrics will produce no more disruption than a familiar melody without lyrics. However, if access to the phonological lexicon is required, a distracter with lyrics will produce a greater disruption than without lyrics. Moreover, if stored separately, spoken lyrics, familiar or unfamiliar, without melody should be less disruptive than familiar melody and lyrics. While the interference-by-content accounts focus on the individuality of each item presented (and their rate of decay, or structure, such as phonological similarity; Baddeley, 1986; Nairne, 1990; Neath, 2000), the perceptual-motor view focuses on sequence factors rather than constituent, individual items, and means of output via motor-planning that enable sequences to be assembled and sub-vocally rehearsed (Hughes & Jones, 2005; Jones et al., 2006, 2007).

3.2 Experiment 1: Method

3.2.1 Participants.

Two hundred and ninety-four participants undertook this experiment (57 male and 237 females; mode 45-49[16]). Participants were recruited from U3A groups, Soroptimist, Rotary, choirs, orchestras, universities, colleges, nurseries, personal contacts, and local community groups following email requests to their activities co-ordinators, poster display, and visits to meetings. Ethical approval for this, and subsequent experiments to be reported, was obtained from the University of Central Lancashire (Appendix A). Community participants received Amazon/Love-to-Shop £5 gift vouchers to reimburse travel expenses, student participants received course credits or chocolate.

[16] Participants ticked an age-range, so no exact ages were obtained.

Since older adults tend to be a more heterogeneous sample than younger adults (e.g., Buczylowska & Petermann, 2016; Mella, Fagot, & de Ribaupierre, 2016) and multiple between-participants conditions were used, to ensure comparability between groups the 150 participants aged over 50 years completed the Addenbrooke's Cognitive Examination Revised (ACE-R) which also gave mini-mental state examination (MMSE) information. Scoring for this experiment ranged between 76-97 (M = 92.75, SD = 4.28) with MMSE[17] scores falling within the normal range expected on the measure between 24 and 30 (M = 28.86, SD 1.19). This indicates positive cognitive functioning for all except two participants over 50 years who scored <82 but scored >24 for MMSE. A later check revealed that the relatively low cognitive scores for these two participants was not necessarily reflected in their recall accuracy, scoring above participants with higher cognitive scores. A mixed analysis of variance (ANOVA) showed Sound Condition × Cognition[18] interaction to be non-significant, $F(2, 296) = 2.042$, $MSE = .224$, $p = .132$, $\eta_p^2 = .014$, supporting an even distribution of participants with comparable cognitive ability across between-participant groups relating to Sound Type[19]. In addition, a visual acuity test (task font 22-24) confirmed acceptable vision, and a test for hearing level, conducted via an audiometer, recorded hearing levels >30dB at 0.5-4.0 kHz for all participants indicated normal hearing[20].

Musical culture.

Prior to beginning the experiment all participants completed a demographic response form which addressed knowledge of "Western musical culture", musical experience, participation, and interest. For Study I a "Western" musical culture was identified by 98.9% of participants, with 2.1% indicating "Other". No participant recorded an African, Asian, or Oriental musical culture.

Musical training and participation.

Participants indicated years of musical training and current participation in a musical activity on their individual demographic response sheet. An ANOVA comparing humming scores of all participants with prior musical training, indicated a rise in accuracy across the three auditory conditions, although non-significant, that corresponded to increased musical training from 0-8+ years. However, performance across all three conditions improved more noticeably following five or more years

[17] A cut off <88 giving 0.94 sensitivity for dementia with <82 giving 0.84. MMSE >24 no cognitive impairment.
[18] Based on ACE-R data.
[19] There was a strong positive correlation between recall scores and cognition (Appendix D)
[20] Normal hearing established at 0-25 dB, some mild hearing loss at whisper level recorded at 20-40 dB.

musical training[21]. In addition, current participation in a choir or instrumental group was a non-significant factor in melody recall performance ability.

3.2.2 Design.

Experiment 1 was conducted as a mixed 3 (Sound Condition) × 4 (Sound Type) design, the dependent variable being accuracy of melody recall via humming. A further mixed ANOVA was conducted to establish prior knowledge of target and distracter melodies through a Recognition test, and an additional 3 (Sound Condition) × 4 (Sound Type) mixed ANOVA examined onset time taken to begin target melody production. The within-participants variable of Sound Condition was classified into three levels, quiet (no distracter), familiar distracter present, and unfamiliar distracter present. The between participants variable of Sound Type was classified into four levels, Familiar Lyrics, Melody, Speech (spoken-lyrics), and Unfamiliar Lyrics, and participants were alternately allocated to each Sound Type (Group) according to the balance of numbers per condition. Each Sound Type had 72 participants with an even distribution of age (except 78 participants participated in the Familiar Lyrics Group). To address repeated measure unsystematic variation, there were six orders of experimental conditions in each Group, twelve participants allocated to each order for each Group.

3.2.3 Materials and apparatus.

In order to capitalize on their familiarity and learning processes, nursery rhymes, traditional poems set to music, were chosen to be target and distracter melodies for this empirical series as they share certain distinctive intrinsic melodic, rhythmic, and lyrical complementary characteristics: a narrative song of stanzaic form; repetitive conjunct rhythm patterns; mostly in a major key; a memorable cantilena; lyrics with a predominantly syllabic setting. The encoding of nursery rhymes encompasses pitch, rhythm, and lyrics, in combination, frequently with gesture (Thomas, 1930), although, their familiarity of structure and predictability can create an ‘earworm’ annoyance (Jakubowski et al., 2016). However, these enduring features contribute to their potential for protracted memory retention of melody and lyrics.

Forty task melodies were taken from collections of popular western culture nursery rhymes, traditional, and patriotic songs. To ascertain familiarity a pre-study questionnaire containing 80 song titles was undertaken by 20 volunteers (none of whom participated in the research experiments). Song titles were rated as “very familiar”, “familiar”, “vaguely familiar”, or “unfamiliar”. If titles were rated “vaguely familiar”

[21] Non-significant interactions: Sound × Training (*p* = .735); Sound × Instrument (*p* = .617); Sound × Choir (*p* = .497).

or “unfamiliar” by more than 55% of volunteers they were discarded, and from the 47 remaining, 40 of the most well-known were selected to be target and distracter melodies (Appendix B).

Musical elements: tempo, tonality, texture, duration, dynamic.

Target and distracter melodies comprised simple duple, triple, quadruple, or compound duple time signatures, a duration range of 12-24 second (*M* = 18.4), and a beat range of 24-48 (*M* = 36.9). All stimuli were computer processed using synthesized sounds of the “Avid Sibelius 7” music software notation program, thereby keeping low level acoustical factors similar across conditions. Tempo of each excerpt was consistent. Pace in western music is indicated by beats per minute (bpm) and the tempo of 120 bpm to a crotchet pulse (quarter note) was used for each excerpt. This produced one beat, and the time between successive beats, in 0.5-second, and corresponds to Allegretto (moderately quick) and March tempi. The human basic pulse rate is 60 bpm and falling musical accents on beats one and three create an actual 60 beat pulse. The bpm system is also used within sequencers to denote tempo so there would be no variance between live recording and computer processed music. To address additional musical elements producing potential confounding variables all melodies had the consistent tonality of C major. This is a commonly used scale in western music as, being derived from the Ionian mode, it uses only natural notes with no accidentals (#s or *b*s). Although not ubiquitous, for most people emotions are influenced by choice of key. C major, representative of the pure emotion of innocence, was considered more appropriate than the passion of B, triumph of D, or sheer joy of E. It also provided a comfortable tessitura (C3-C5 octave span) for the vocalist. Each of 30 target melodies ended on the tonic and began with a note of the tonic triad, 12 with tonic, 13 with dominant (three as anacruses), and five with mediant (one as anacrusis). All task melodies had a monophonic texture with legato phrasing and instrumental conditions were processed using a concert flute in C tone timbre. Being a treble clef, reed-less, non-transposing instrument with a light, bright, clear penetrating sound, this was considered an appropriate medium within Sibelius to convey a smooth, poetic solo line befitting the nursery rhyme task. Time values comprised quavers (eighth notes) or longer duration (crotchets and minims) with an occasional dotted quaver + semiquaver combination. The dynamic consistently mezzo-forte *(mf)* with no tone gradation.

Matching and mismatching.

For unfamiliar matched melodies, while tonality, rhythmic pattern, compass, harmonic framework, tempo, dynamic, and implied harmony using basic diatonic chords of tonic, dominant, and subdominant (I, V, IV), remained consistent to the original melody, the pitch contour was altered to create novel interval progressions whilst retaining C as the tonic. All ended (analogous to their respective original versions) with an implied perfect/plagal cadence to create finality. Unlike the unfamiliar lyrics from obscure songs previously used in studies (Cuddy et al., 2012) here novel lyrics were specifically created. They comprised a predominantly syllabic setting with limited melisma, similar repetitive metre, and general syntactic structure of the original lyrics. (Examples of auditory distracter conditions for the nursery rhyme "London Bridge" shown in Figure 1, a-h.)

Figure 1a. Melody Group (familiar melody) auditory distracter condition.

b. Melody Group (unfamiliar melody) auditory distracter condition.

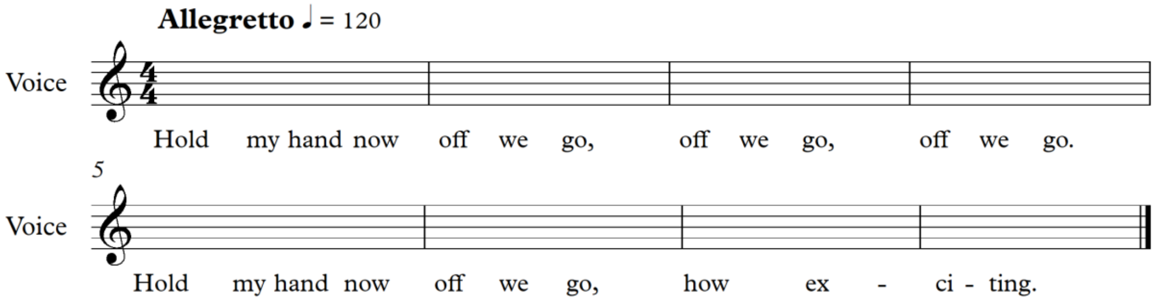
c. Familiar Lyrics Group (familiar melody) auditory distracter condition.

d. Familiar Lyrics Group (unfamiliar melody) auditory distracter condition.

e. Unfamiliar Lyrics Group (familiar melody) auditory distracter condition.

f. Unfamiliar Lyrics Group (unfamiliar melody) auditory distracter condition.

g. Speech Group (familiar spoken-lyrics) auditory distracter condition.



h. Speech Group (unfamiliar spoken-lyrics) auditory distracter condition.

Performance.

A female contralto voice, a trained musician, produced a live a cappella performance into a laptop computer via a line-in USB microphone, using the sound recording and editing-program Audacity. To control for uniformity between sung-lyrics and spoken-lyrics, the same voice recorded familiar spoken-lyrics and matched spoken-lyrics on a monotone with limited vocal inflection. Speech was paced to the tempo of the sung version in keeping with previous research (Simmons-Stern et al., 2012). In this way sung and spoken tracks were of equal duration.

Stimuli were looped from their original times to run for 32-second to allow equal duration with the 32-second time requirements to complete humming of each target melody. The AUP project files created were then exported from Audacity to create .WAV files which were subsequently played through media player to develop the task using E-Prime 2.0 Psychology software tool. Four song-set orders (A, B, C, D) from the 40 chosen melodies were created, each set of ten having the same or as near same humming length of 184-185 seconds (Appendix C). A random order of song titles via E-Prime was preferred to a fixed order.

3.2.4 Procedure.

To assess whether melody and lyrics are integrated or represented independently various combinations of melody and lyrics were manipulated and delivered to predict if retrieval was limited or affected through suffering impairment due to the presence of a distracter. All testing was undertaken individually in a quiet room in the presence of the researcher. The laptop computer internal microphone with a 20dB boost was used for recording and all participants wore a set of Sennheiser HD201 headphones, attached to the computer, through which all irrelevant audio was presented[22]. To ensure standardization of instructions to participants tasks were presented using E-Prime

[22] Irrelevant sounds averaged 62dB (A) as measured by a sound level meter attached to the headphones.

program (2). Instructions to “Press the spacebar when you are ready to progress” accompanied by an audible bleep were consistent for each movement through the program.

Following two practice trials in quiet, all participants experienced two distracter sound condition blocks (familiar and unfamiliar), and a no-sound block (quiet). The order was counterbalanced to reduce demand characteristics and song-titles within each block randomly determined. This enabled 78 participants in the Familiar Lyrics Group to experience familiar melody (familiar sung-lyrics) and unfamiliar matched melody (familiar sung-lyrics), 72 participants in the Melody Group to experience familiar melody (no lyrics) and unfamiliar matched melody (no lyrics), 72 in the Speech Group experienced familiar spoken-lyrics and unfamiliar spoken-lyrics, with 72 in the Unfamiliar Lyrics Group having familiar melody (unfamiliar sung-lyrics) and unfamiliar matched melody (unfamiliar sung-lyrics). For example, a participant in the Melody Group was required to hum ten target melodies, in each of three condition blocks, according to the title visible on the computer screen, while accompanied by to-be-ignored sound comprising either a familiar melody, an unfamiliar melody, or no sound (quiet). Target melody titles were taken from song-set A or B or C depending on the program melody set order in the six orders of distraction. Both familiar and unfamiliar matched distracter melodies were taken from melody set D (see Appendix C). Participants in the Familiar Lyrics Group, Speech Group, and Unfamiliar Lyrics Group followed the same procedure with their respectively different distracters. The visual appearance of the target melody title and the distracter sound were set to begin simultaneously. Finally, after participants had undertaken all experimental trials, a Recognition test presented all 40 melodies (including song-set D) in order to control for prior knowledge. Participants were instructed to press the “Y” button if familiar, or “N” if unfamiliar, according to the name of each target melody. If “Y” was pressed they were instructed to hum the melody again, in quiet condition. If “N” was pressed further humming was not required. All humming attempts from each experimental trial was recorded through the computer and converted by the E-prime program into individual .WAV files.

3.2.5 Assessment.

Unlike recall assessment procedures used by Sloboda and Parker (1985) audio recordings from participants’ melody recall for Study I was not transcribed into musical notation by the author/assessor or raters. Assessment was undertaken direct from assessor/rater judgement when listening to the individual recording of participants’ audio files. Although our unique voice is the musical instrument that we all possess, judgement as to the quality of an individual's accuracy of pitch production is often based on aesthetical criteria (Larrouy-Maestri & Morsomme, 2014). The assessment of participants' performance needed an objective acoustic method of assessment with an understood definition of melodic accuracy, and strategies consistently applied. Errors in performance were considered probable with melodic contour, in interval movement, or with intonation. However, it has been shown that as listeners we award a high level of vocal generosity to singers over instrumentalists when judging notes to be out-of-tune (Hutchins, Roquet, & Peretz, 2012) or accurate as compared to precise (Pfordresher, Brown, Meier, Belyk, & Liotti, 2010). This "vocal generosity", understood by the researcher/assessors, was adopted into performance assessment procedure for the current study. For Experiment I only accuracy of melodic contour, to show recognition of the target melody, was assessed, a broad tolerance awarded to misplaced intervals and intonation. This allowed for a small interval error—such as a major 3rd hummed as a perfect 4th—to be assessed positively providing the contour of the proceeding bar(s) maintained a correct trajectory from the point of the misplaced interval. If the contour were incorrectly followed from the point of deviation then the bar(s) would be deemed incorrect.

Recordings of raw data .WAV files from all vocal performances of retrieved target melodies were measured by the author/assessor using a scale ranging from zero, inaccurate, to four, accurate and fluent, based on established musical performance assessment criteria used by exam boards (e.g., Edexcel), known and understood by the author/assessor. The scale was designed to allow a zero score for no rewardable material, the total number of bars for each melody then being divided into four sections of equal number to match the criterion descriptors, see Table 3.1.

Table 3.1

Assessment criteria for the measurement of humming retrieval accuracy

Score	0	1	2	3	4
Criteria	inaccurate, no rewardable content	limited accuracy, repeated errors	broadly accurate, some errors	mainly accurate	accurate + fluent
total bars	bars correct	bars correct	bars correct	bars correct	bars correct
8	0	1-2	3-4	5-6	7-8
12	0	1-3	4-6	7-9	10-12
14	0	1-3.5	3.5-7	7-10.5	10.5-14
16-15	0	1-4	5-8	9-12	13-16
20	0	1-5	6-10	11-15	16-20

This criterion provided validity for measure of exact purpose, pitch accuracy, without reference to rhythm or other musical elements. In addition, participants' .WAV files were burnt to CD for two independent, musically trained raters (music teachers and orchestral performers) to assess. They compared performances from half of each Sound Type Group for accuracy to the target melody following the same scoring criterion. To assess inter-rater reliability an intraclass correlation coefficient (ICC) was conducted and showed a high consistency between researcher and independent raters[23], Cronbach's $\alpha = .966$. Furthermore, participants' vocal performances were judged as recognizable by non-musically trained raters (secondary school teachers). Here, only recognition of the target melody was required. An ICC test between the two non-musically trained raters again showed a high consistency, Cronbach's $\alpha = .988$.

Onset time, in seconds, measuring time taken from visual presentation of the target title, to search retrieval of the target melody from participants' memory, and begin recall production via humming of the target melody was also recorded as an additional indicator of response across the three auditory conditions.

Here, and applicable elsewhere in this empirical series of study, all data were analysed using the analysis of variance (ANOVA) technique with an alpha level of .05. For each main effect and interaction effect, we reported the classical *F* and *p* values along with an estimate of the effect size (η_p^2). In addition to controlling error rate for multiple tests a balance between Type I error and statistical power needed to be maintained in order to reduce the probability of increased Type II error. Although Holm-Bonferroni does control for Type I error (and is more powerful than Bonferroni), as ANOVA was chosen to avoid multiple t-tests, and LSD has been applied in prior distraction studies using ANOVA (e.g., Marsh et al., 2009), it was considered an appropriate post-hoc test to identify significant differences in this empirical series.

3.3 Results

The main purpose of Study I was to identify the effect on accuracy of melody retrieval performance by humming, in the presence or absence of a familiar or unfamiliar sound distracter, from 294 participants by comparing means for each of three levels of Sound variable, quiet, familiar, and unfamiliar. Each humming sound file from each participant was listened to and assessed by the researcher/raters according to criteria identified in Table 3.1 above.

---

[23] Any discrepancies greater than 2 were assessed again for reconsideration.

Two participants aborted the E-prime program (by pressing an irrelevant computer key part way through the final Recognition test), but their auditory conditions data was included in the analysis. A 3 (Sound Condition: quiet, familiar, unfamiliar) × 4 (Sound Type: Familiar Lyrics, Melody, Speech [spoken-lyrics], Unfamiliar Lyrics) mixed analysis of variance (ANOVA) was conducted**.** Figure 2 shows that baseline data was relatively consistent across Sound Type in the quiet condition (although slightly higher for the Unfamiliar Lyrics Group in comparison to the Melody Group) but decreased in the presence of an unfamiliar distracter and decreased further in the presence of a familiar distracter for all Sound Type Groups. Accuracy of recall performance appeared to be affected by both the auditory condition of the distracter presented, instrumental melody, spoken-lyrics, or sung-lyrics, and according to the familiarity of the irrelevant condition, the impact of a familiar distracter being greater than that of an unfamiliar distracter across all Sound Type Groups.

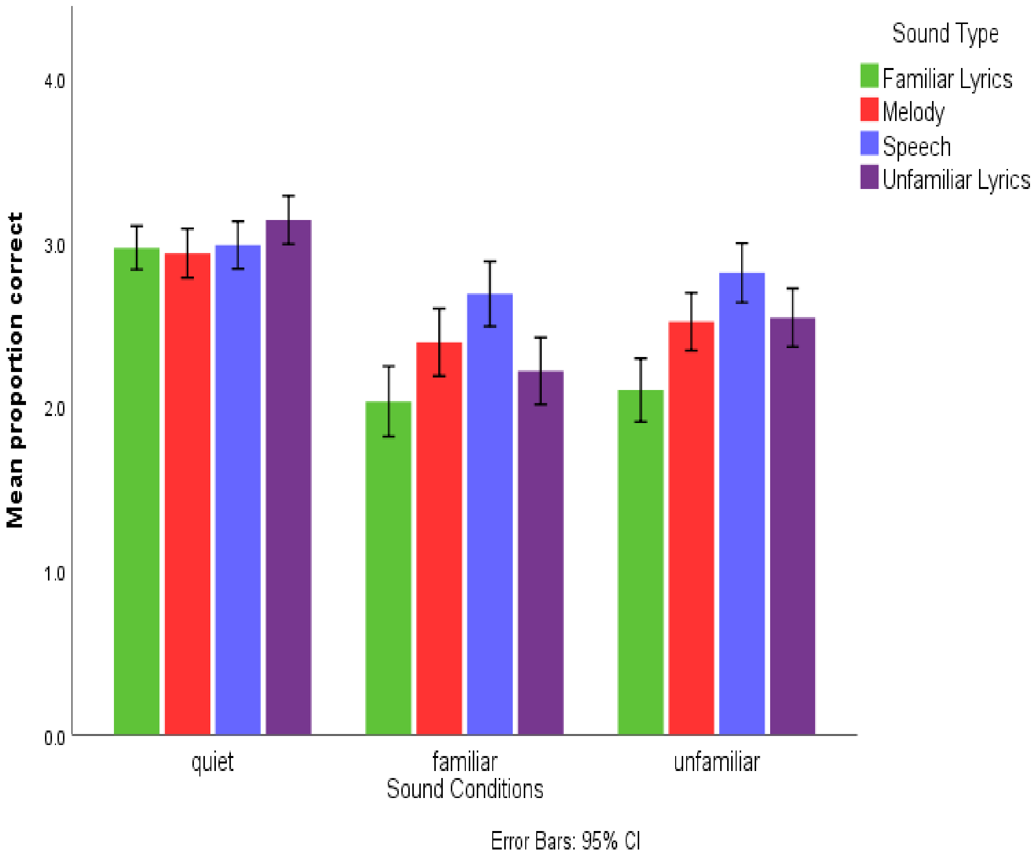


Figure 2. Mean proportion retrieval accuracy for melody humming across the quiet, familiar, and unfamiliar Sound Conditions according to the Sound Type Group variable.

Results from a mixed ANOVA[24] showed a significant main effect of Sound Condition, $F(1.960, 568.49) = 161.226$, $MSE = .230$, $p < .001$, $\eta^2_p = .357$, indicating recall performance was affected by the presence of both familiar and unfamiliar task-irrelevant sound. There was also a significant between-participants main effect of Sound Type, $F(3, 290) = 6.089$, $MSE = 1.335$, $p < .001$, $\eta^2_p = .059$. In addition, there was a significant interaction between Sound Condition × Sound Type, $F(5.881, 568.494) = 10.686$, $MSE = .230$, $p < .001$, $\eta^2_p = .100$, indicating that the disruptive potency of Sound Condition (familiar, unfamiliar) on melody recall performance was influenced by Sound Type (e.g., Familiar Lyrics, Melody, Speech, Unfamiliar Lyrics). A follow-up simple effects analysis was undertaken to decompose the Sound Condition × Sound Type interaction.

The first wave of analysis compared means for Sound Conditions (quiet, familiar, unfamiliar) as a function of Sound Type (Familiar Lyrics, Melody, Speech, Unfamiliar Lyrics). Initial comparisons were focused on determining whether the presence of sound, regardless of its familiarity, determined disruption by comparing quiet and unfamiliar sound conditions within each sound type. Performance was poorer in the unfamiliar condition compared to quiet condition for Familiar Lyrics ($MD = .868$, $SE = .069$, $p < .001$; 95% CI [.733, 1.003]), Melody ($MD = .418$, $SE = .071$, $p < .001$; 95% CI [.278, .559]), Speech ($MD = .169$, $SE = .071$, $p = .018$; 95% CI [.029, .310]), and Unfamiliar Lyrics ($MD = .596$, $SE = .071$, $p < .001$; 95% CI [.455, .736]) Groups.

This demonstrated that accuracy of humming performance was poorer when accompanied by sound regardless of whether it comprised non-speech (melody) or speech (spoken-lyrics) and regardless of whether it comprised a familiar component (e.g., familiar lyrics) or not (e.g., unfamiliar lyrics). Secondary comparisons investigated whether familiarity effects occurred within each group by comparing familiar and unfamiliar Sound Conditions. For the Unfamiliar Lyrics Group performance was significantly poorer in the familiar compared with unfamiliar Sound Condition ($MD = -.325$, $SE = .081$, $p < .001$; 95% CI [-.485, -.165]), thereby demonstrating a familiarity effect. See Table 3.2 for all familiar *versus* unfamiliar pairwise comparisons.

[24] Huynh-Feldt reported following sphericity violation $\varepsilon = .980$

Table 3.2

Sound Condition comparisons: familiar melody (fam) versus unfamiliar melody (unf) within Sound Type Group

Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Familiar Lyrics	fam *vs.* unf	-.069	.078	.377	-.223	.085
Melody	fam *vs.* unf	-.125	.081	.126	-.285	.035
Speech	fam *vs.* unf	-.128	.081	.117	-.288	.032
Unfamiliar Lyrics	fam *vs.* unf	-.325	.081	< .001	-.485	-.165

Note. 95% Confidence Intervals

However, although there was a trend for two conditions, no such familiarity effect emerged for Melody (p = .126), Speech (p = .117), or Familiar Lyrics (p = .377) Groups. Thus, production of a target melody (e.g., 'Mary, Mary, quite contrary') was impaired by exposure to the combination of an unfamiliar melody (e.g., 'Hold my hand) with familiar lyrics (e.g., 'London Bridge' see Figure 1d) as compared with an unfamiliar melody combined with unfamiliar lyrics (e.g., 'Hold my hand'; see Figure 1f). However, the presence of familiar lyrics accompanied by familiar melody (e.g., 'London Bridge'; see Figure 1c) produces no more disruption to production of a target melody (e.g., 'Mary, Mary quite contrary') than familiar lyrics (e.g., 'London Bridge') accompanied by unfamiliar melody (e.g., 'Hold my hand; Figure 1d). This suggests that familiarity—either provided by melody or lyrics—within a to-be-ignored song produces disruption relative to when familiarity is not present in neither melody nor lyrics.

The second wave of analysis explored differences across Sound Type Groups as a function of Sound Condition. As indicated earlier, there was a marginally significant difference between baseline (quiet) conditions of the Melody Group and Unfamiliar Lyrics Group (MD = -.203, SE = .103, p = .05; 95% CI [-.405, .000]), with slightly superior performance for the Unfamiliar Lyrics Group. All shown in Table 3.3.

In the context of unfamiliar sound, humming performance was poorer when accompanied by Melody as compared with Speech (MD = -.299, SE = .131, p = .023; 95% CI [-.577, -.041]). Thus, humming accuracy performance was poorer when the target melody (e.g., 'Mary, Mary, quite contrary') was attempted in the presence of unfamiliar melody (e.g., 'Hold my hand'; see Figure 1b) compared to spoken-lyrics (e.g., 'Hold my hand'; see Figure 1h).

Table 3.3

Sound Condition comparison across Sound Type Group

Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Melody *vs.* Unfamiliar Lyrics	quiet	-.203	.103	.050	-.405	.000
Familiar Lyrics *vs.* Melody	familiar	-.361	.145	.013*	-.646	-.076
Familiar Lyrics *vs.* Speech		-.657	.145	< .001*	-.942	-.372
Melody *vs.* Speech		-.296	.148	.046*	-.587	-.005
Speech *vs.* Unfamiliar Lyrics		.471	.148	.002*	.180	.762
Unfamiliar Lyrics *vs.* Familiar Lyrics		.186	.145	.200	-.099	.471
Melody *vs.* Unfamiliar Lyrics		.175	.148	.237	-.116	.466
Familiar Lyrics *vs.* Melody	unfamiliar	-.417	.129	.001*	-.670	-.164
Familiar Lyrics *vs.* Speech		-.715	.129	< .001*	-.968	-.462
Familiar *vs.* Unfamiliar Lyrics		-.442	.129	.001*	-.695	-.189
Melody *vs.* Speech		-.299	.131	.023*	-.557	-.041
Speech *vs.* Unfamiliar Lyrics		.274	.131	.038*	.016	.532
Unfamiliar Lyrics *vs.* Melody		.025	.131	.849	-.233	.283

Note. * denotes significance.

In line with the notion that one component of familiarity within song may be sufficient to produce disruption of melody retrieval, performance was poorer when unfamiliar melody combined with familiar lyrics compared to unfamiliar melody alone (*MD* = -.417, *SE* = .129, *p* < .001; 95% CI [-.164, -.670]) unfamiliar speech (*MD* = -.715, *SE* = .129, *p* < .001; 95% CI [-.968, -.462]), and crucially unfamiliar lyrics (*MD* = -.442, *SE* = .129, *p* = .001; 95% CI [-.695, -.189]). Thus, humming accuracy was poorer when the target melody (e.g., 'Mary, Mary, quite contrary') was attempted in the presence of an unfamiliar melody (e.g., 'Hold my hand') combined with familiar lyrics (e.g., 'London Bridge'; Figure 1d) compared to unfamiliar melody alone (e.g., 'Hold my hand'; Figure 1b) or unfamiliar spoken-lyrics (e.g., 'Hold my hand'; Figure 1h). The combination of unfamiliar melody and lyrics, however, was more disruptive than unfamiliar speech (*MD* = -.274, *SE* = .131, *p* = .038; 95% CI [-.532, -.016]), perhaps owing to the presence of melody within the sequence (since unfamiliar melody produces greater disruption than unfamiliar spoken-lyrics, see earlier).

In relation to familiar sound, melody was more disruptive than speech (*MD* = -.296, *SE* = .148, *p* = .046; 95% CI [-.116, .466]). Thus, when the melody was familiar (e.g., ‘London Bridge’; Figure 1a) humming accuracy was more impaired than when familiar lyrics were spoken (e.g., ‘London Bridge’; Figure 1g). The combination of familiar melody and familiar lyrics produced greater disruption to melody recall accuracy than familiar melody alone (*MD* = -.361, *SE* = .145, *p* = .013; 95% CI [.076, .646]). Furthermore, the combinations of familiar melody with familiar lyrics and familiar melody with unfamiliar lyrics produced more disruption than familiar spoken-lyrics alone (*MD* = -.657, *SE* = .145, *p* < .001; 95% CI [-.942, -.372], and (*MD* = -.471, *SE* = .148, *p* = .002; 95% CI [-.762, -.180] respectively). Therefore, production of a target melody (e.g., ‘Mary, Mary, quite contrary’) was impeded more by combination of familiar melody and familiar lyrics (e.g., ‘London Bridge’; Figure 1c) and combination of familiar melody (e.g., ‘London Bridge’) with unfamiliar lyrics (e.g., ‘Hold my hand now off we go’; Figure 1e) than familiar melody by itself (e.g., ‘London Bridge’; Figure 1a). The combination of familiar melody with familiar lyrics was no more disruptive than the combination of familiar melody with unfamiliar lyrics (*p* = .200). However, somewhat surprisingly, performance with unfamiliar lyrics accompanied by familiar melody was not lower than performance in the familiar melody only condition (*p* = .237). Thus, target melody production (e.g., Mary, Mary, quite contrary) was no more impaired by concurrently ignoring a combination of unfamiliar lyrics (e.g., ‘Hold my hand’) with familiar melody (e.g., ‘London Bridge’; Figure 1e) than from ignoring a familiar melody alone (e.g., ‘London Bridge’; Figure 1a)

To summarize, foregoing results showed retrieval of melody—as indexed by accuracy of humming performance—was adversely affected by all Sound Conditions as compared to quiet. Although there was a tendency for familiar sound to be more disruptive than unfamiliar sound in two of the four groups (Melody, Speech), a significant difference only arose for the Unfamiliar Lyrics Group wherein unfamiliar lyrics combined with familiar melody was more disruptive than unfamiliar lyrics combined with unfamiliar melody, suggesting incongruity in familiarity between lyrics and melody is important in determining disruption. Absence of a familiarity effect within the Familiar Lyrics Group may be because familiarity on one of the two dimensions (lyrics or melody) suffices to produce disruption. Melody was generally more disruptive than Speech (spoken-lyrics) regardless of its familiarity. While the combination of unfamiliar melody with familiar lyrics was clearly more disruptive than unfamiliar melody, spoken-lyrics, and unfamiliar melody combined with unfamiliar lyrics, differences in relation to pairings within the context of familiar sounds were less clear-cut. While the combination of familiar melody with lyrics impaired recall accuracy more than familiar melody and familiar spoken-lyrics, it was not more disruptive than familiar melody combined with unfamiliar lyrics and this condition, although more disruptive than familiar spoken-lyrics, was no more disruptive than familiar melody.

3.3.1 Recognition test.

Following distracter trials a Recognition test, to control for prior knowledge of target and distracter melodies, was conducted in quiet conditions. All 40 nursery rhyme titles (including distracter rhymes) were presented again for melody recall. Participants were instructed to press the "Y" key on the computer keyboard to indicate that the melody of the target rhyme was "familiar" to them and they were then instructed to hum the melody again. The total mean of "familiar" melodies acknowledged was 26.85 (*SD* = 9.74) indicating 67% knowledge of target and distracter melodies. The number of known melodies recognized and attempted, from a possible 10 for each trial, was 8.22 (*SD* = 1.87) in the quiet condition, 7.63 (*SD* = 2.23) in the familiar condition, and 7.84 (*SD* = 2.13) in the unfamiliar condition, generating an 79% total mean of known recognized melodies. It should be noted however, that in the Recognition test there was some reticence to commit to a "Y" if only part of the melody was known, whereas in the trials partial retrieval of melody had been recorded. This discrepancy resulted in a slightly lower total percentage in the Recognition test than trials would indicate.

A mixed ANOVA was then run for mean recall scores from the two sung-lyric Sound Type Group distracter trials, the lowest scoring, against mean recall scores from the final Recognition test in the absence of a distracter, with Sound Type Group as the between-participants factor, and this showed increased accuracy against quiet (see Table 3.4) indicating differences in the trials resulting from the distracter Sound Condition rather than lack of knowledge of the target melodies.

Table 3.4

Mean retrieval scores for familiar and unfamiliar Sound Conditions, with and without distracter, for Familiar Lyrics and Unfamiliar Lyrics Sound Type Groups

Sound Type Group	familiar distracter	Recognition without distracter	unfamiliar distracter	Recognition without distracter
Familiar Lyrics	2.03 *(.95)*	3.07 *(1.08)*	2.10 *(.86)*	3.01 *(1.10)*
Unfamiliar Lyrics	2.22 *(.87)*	3.31 *(.61)*	2.54 *(.76)*	3.34 *(.63)*

Results[25] showed a significant main effect of Sound Condition, $F(2.347, 680.707) = 216.046$, $MSE = .376$, $p < .001$, $\eta_p^2 = .427$, suggesting increased accuracy due to absence of a distracter rather than lack of prior knowledge. There was a significant between-participant effect of Sound Type, $F(3, 290) = 6.004$, $MSE = 1.805$, $p = .001$, $\eta_p^2 = .058$, indicating some difference in accuracy for the Sound Type Groups. A simple effects analysis revealed significance between all the Familiar Lyric Group pairings see Table 3.5.

Table 3.5

Pairwise comparison for Sound Type in Recognition Test

Sound Type Group	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Familiar Lyrics *vs.* Melody	-.282	.110	.011	-.498	-.066
Familiar Lyrics *vs.* Speech	-.455	.110	< .001	-.671	-.239
Familiar Lyrics *vs.* Unfamiliar Lyrics	-.298	.111	.007	-.514	-.081

As expected, pairwise comparisons for Sound Conditions revealed significance (all *p*'s < .001) between, all familiar and unfamiliar Sound Conditions except between the two Recognition test scores themselves ($MD = -.038$, $SE = .032$, $p = .226$; 95% CI [-.100, .024]), identifying increased scoring in the quiet Recognition phase in comparison with scoring in the distracter trials from the two lowest scoring lyric groups. In addition, there was a significant Sound Condition × Sound Type interaction, $F(7.042, 680.707) = 4.495$, $MSE = .376$, $p < .001$, $\eta_p^2 = .044$. A simple effects analysis for Sound Condition as a function of Sound Type identified this effect to be driven by lyrics groups, but particularly by the sung-lyrics groups see Table 3.6.

[25] Huynh-Feldt reported following sphericity violation $\varepsilon = .782$

Table 3.6

Pairwise comparison for Sound Condition across Sound Type between familiar and unfamiliar Sound Condition trial scores and Recognition (Recog) test scores

Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Familiar Lyrics *vs.* Melody	familiar	-.361	.145	.013*	-.646	-.076
Familiar Lyrics *vs.* Speech		-.657	.145	< .001*	-.942	-.372
Melody *vs.* Speech		-.296	.148	.046*	-.587	-.005
Speech *vs.* Unfamiliar Lyrics		.471	.148	.002*	.180	.762
Unfamiliar Lyrics *vs.* Melody		-.175	.148	.237	-.466	.166
Familiar Lyrics *vs.* Melody	unfamiliar	-.417	.129	.001*	-.670	-.164
Familiar Lyrics *vs.* Speech		-.715	.129	< .001*	-.968	-.462
Familiar Lyrics *vs.* Unfamiliar Lyrics		-.442	.129	.001*	-.695	-.189
Melody *vs.* Speech		-.299	.131	.023*	-.557	-.041
Unfamiliar Lyrics *vs.* Melody		.025	.131	.849	-.233	.283
Familiar Lyrics *vs.* Speech	Recog unfamiliar	-.312	.129	.016*	-.567	-.058
Familiar Lyrics *vs.* Unfamiliar Lyrics		-.321	.129	.014*	-.575	-.066

Note. * denotes significance.

3.3.2 Onset time.

For each participant, the onset time (OST) taken to begin production of the target melody was recorded. This was similar between the Sound Type Groups when the distracter was absent, but interestingly there was less similarity when a distracter was present. OST rose in each group when presented with the familiar distracter, but more noticeably for the Familiar Lyrics Group see Figure 3.

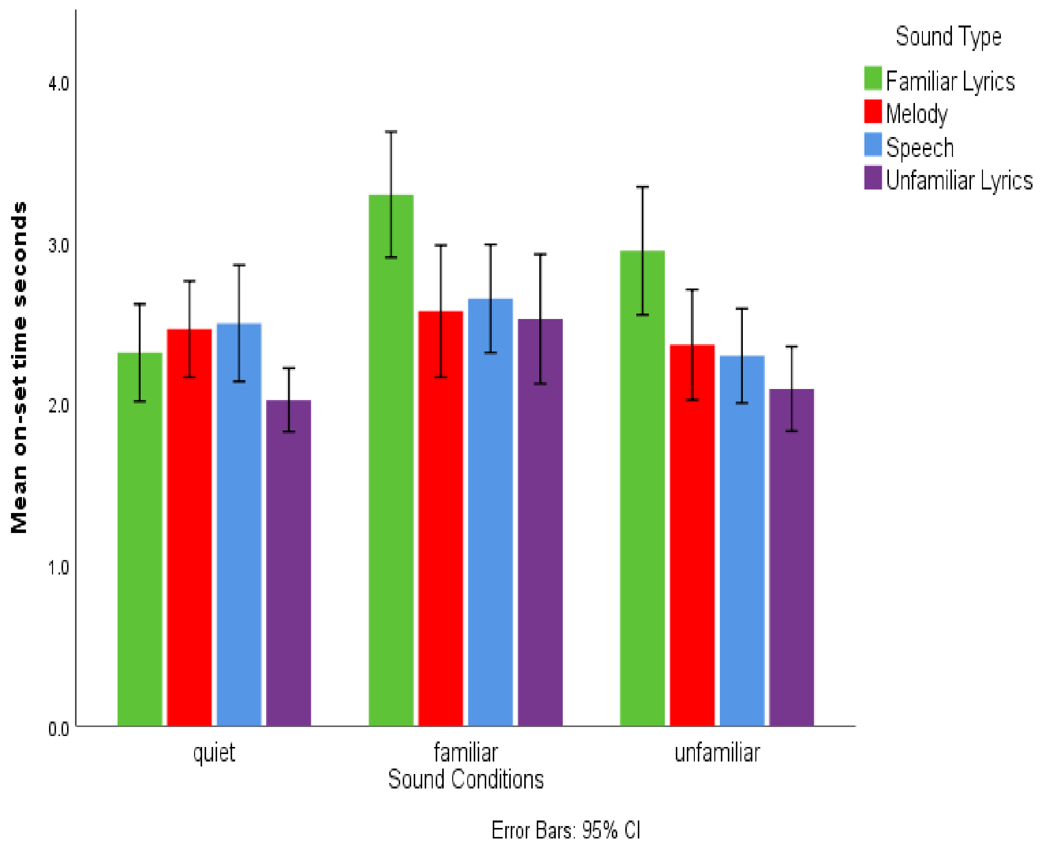


Figure 3. Mean onset time to begin humming the target melody across Sound Conditions by Sound Type Group.

Results from a 3 (Sound Condition: quiet, familiar, unfamiliar) × 4 (Sound Type: Familiar Lyrics, Melody, Speech, Unfamiliar Lyrics) mixed ANOVA[26] showed a significant main effect of Sound, $F(1.942, 563.235) = 12.587$, $MSE = 1.259$, $p < .001$, $\eta_p^2 = .042$. The between-participants effect of Sound Type was also significant, $F(3, 290) = 3.981$, $MSE = 3.955$, $p = .008$, $\eta_p^2 = .040$. In addition, there was a significant Sound Condition × Sound Type interaction, $F(5.827, 563.235) = 3.185$, $MSE = 1.259$, $p = .005$, $\eta_p^2 = .032$, that showed on-set time varied according to Sound Type.

A follow-up simple effects analysis was performed to unpack the significant interaction between Sound Condition and Sound Type. In the first wave of analysis the means for Sound Conditions (quiet, familiar, unfamiliar) as a function of Sound Type (Familiar Lyrics, Melody, Speech, Unfamiliar Lyrics) were compared. Initial comparisons were focused on determining whether the mere presence of sound, regardless of its familiarity, slowed onset time by comparing quiet and unfamiliar sound

[26] Huynh-Feldt reported following sphericity violation $\varepsilon = .971$

conditions within each Sound Type. Onset time was slower for unfamiliar as compared to quiet for the Familiar Lyrics Group (*MD* = -.632, *SE* = .158, $p < .001$; 95% CI [-.943, -.322]), but not for the Melody Group ($p = .560$), Speech Group ($p = .221$), or Unfamiliar Lyrics Group ($p = .679$). Subsequently, whether familiarity effects occurred within each group was investigated by comparing unfamiliar and familiar sound conditions. A familiarity effect was observed for the Unfamiliar Lyrics Group (*MD* = -.435, *SE* = .188, $p = .022$; 95% CI [-.805, -.064]) and the Familiar Lyrics Group (*MD* = -.349, *SE* = .181, $p = .055$; 95% CI [- .705, .007]) but not for the Melody Group ($p = .273$) or Speech Group ($p = .060$), although the latter comparison approached significance. While analysis of humming melody accuracy demonstrated significant differences between all familiar Sound Conditions and quiet, for onset time this difference only arose for the Familiar Lyrics Group (*MD* = -.981, *SE* = .191, $p < .001$; 95% CI [-1.356, -.605]) and the Unfamiliar Lyrics Group (*MD* = -.503, *SE* = .199, $p = .012$; 95% CI [-.894, -.112]; Melody, $p = .576$; Speech, $p = .438$).

These results demonstrates that onset of humming was delayed by the presence of familiar lyrics combined with unfamiliar melody (e.g., ‘London Bridge’; Figure 1d) compared to quiet. Moreover, the combination of familiar lyrics with familiar melody (e.g., ‘London Bridge’; Figure 1c) delayed humming onset relative to the combination of familiar lyrics (e.g., ‘London Bridge’) with unfamiliar melody (e.g., ‘Hold my hand’; Figure 1d). There was also a tendency for familiar spoken-lyrics (e.g., ‘London Bridge’; Figure 1g) to delay humming onset more than unfamiliar spoken-lyrics (e.g., ‘Hold my hand’; Figure 1h), and for unfamiliar lyrics (e.g., ‘Hold my hand’) accompanied by familiar melody (e.g., ‘London Bridge’; Figure 1e) to produce more disruption than unfamiliar lyrics combined with unfamiliar melody (e.g., ‘Hold my hand’; Figure 1f).

A second wave of analysis explored differences across Sound Type Group as a function of Sound Condition. This analysis revealed some differences at baseline, quiet condition with participants in the Unfamiliar Lyrics Group who commenced humming faster than participants in the Melody Group (*MD* = .439, *SE* = .212, $p = .039$; 95% CI [.022, .856]) and Speech Group (*MD* = .475, *SE* = .212, $p = .026$; 95% CI [.058, .892]). For unfamiliar sound onset time was slower for Familiar Lyrics Group compared with Melody Group (*MD* = -.582, *SE* = .233, $p = .013$; 95% CI [-1.041, -.124]), Speech Group (*MD* = -.652, *SE* = .233, $p = .005$; 95% CI [-1.110, -.193]), and Unfamiliar Lyrics Group (*MD* = -.857, *SE* = .233, $p < .001$; 95% CI [-1.316, -.399]) see Table 3.7.

Table 3.7

OST across Sound Type as a function of Sound Condition

Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Melody *vs*. Unfamiliar Lyrics	quiet	.439	.212	.039	.022	.856
Speech *vs*. Unfamiliar Lyrics		.475	.212	.026	.058	.892
Familiar Lyrics *vs*. Melody	familiar	.724	.271	.008	.190	1.257
Familiar Lyrics *vs*. Speech		.645	.271	.018	.111	1.178
Familiar *vs*. Unfamiliar Lyrics		.771	.271	.005	.238	1.305
Familiar Lyrics *vs*. Melody	unfamiliar	.582	.233	.013	.124	1.041
Familiar Lyrics *vs*. Speech		.652	.233	.005	.193	1.110
Familiar *vs*. Unfamiliar Lyrics		.857	.233	< .001	.399	1.316

Thus, the combination of familiar lyrics (e.g., ‘London Bridge’) with an unfamiliar melody (e.g., ‘Hold my hand’; Figure 1d) impaired melody retrieval time relative to the presence of unfamiliar melody (e.g., ‘Hold my hand’; Figure 1b), unfamiliar spoken-lyrics (e.g., ‘Hold my hand’; Figure 1h) and unfamiliar melody combined with unfamiliar lyrics (e.g., Hold my hand’; Figure 1f).

To summarize, results for OST showed that melody retrieval time was not impacted by all unfamiliar sound conditions. Indeed, only unfamiliar melody combined with familiar lyrics impaired performance relative to quiet. Further, familiarity effects were only observed for the Familiar Lyrics Group (familiar melody with familiar lyrics against unfamiliar melody with familiar lyrics) and the Unfamiliar Lyrics Group (familiar melody with unfamiliar lyrics against unfamiliar melody with unfamiliar lyrics). Moreover, onset of melody retrieval was delayed by the presence of familiar lyrics combined with unfamiliar melody compared with unfamiliar melody, unfamiliar spoken-lyrics, and unfamiliar melody combined with unfamiliar lyrics. Finally, melody retrieval time was impaired by familiar lyrics combined with familiar melody compared with familiar melody, familiar spoken-lyrics, and unfamiliar lyrics combined with familiar lyrics.

3.4 Discussion

The idea that an act of music performance can be influenced, perhaps detrimentally, by concurrent presence of a different musical performance is intuitive. Central to this current study was the need to perform a familiar melody. Results clearly demonstrate an increased disruption to the vocal-motor system generated from competing irrelevant melody/lyrics well-formed motor plans during target production. Hitherto, concentrated efforts on testing the notion of embodied cognition whereby it is proposed that musical performance during concurrent irrelevant musical sound, as compared to recognition or identification of musical patterns/notes, demands an interplay of musical perception, motor-planning and action have not been undertaken (Leman, 2007, Leman & Maes, 2014). A viewpoint based on the perceptual-gestural view (Hughes & Jones, 2005; Jones et al., 2006, 2007) is that motor-systems required to produce the target performance are hampered by the presence of an irrelevant musical performance, more so when irrelevant music is familiar to the performer. It is also noted that any disruption of musical performance might be understood as interference deriving from within the cognitive architecture (Peretz & Coltheart, 2003; Roediger, Gallo, & Geraci, 2002).

As compared to quiet, irrelevant music would seem to have a disruptive impact on short-term memory tasks for visual items (e.g., Cassidy & MacDonald, 2007; Nittono, 1997). However, a systematic evaluation of music components responsible for disruption has not been undertaken. In the context of short-term memory recall focus has been on melody, and manipulations have not routinely included familiar and unfamiliar music (Alley & Greene, 2008; Silverman, 2007), sung vocals (Nittono, 1997; Pring & Walker, 1994), a spoken condition (Iwanaga & Ito, 2002), or a quiet condition to establish a baseline (Iwanaga & Ito, 2002; Pring & Walker, 1994). The nature of disruption of short-term memory recall by task-irrelevant music is likely to be qualitatively distinct from that obtained from paradigms for which melody is to be retrieved. (This issue is explored in Study III.)

3.5 Comparison of overall results with previous findings

The focus of the current study was to differentiate between familiar and unfamiliar melody and lyrics, and to compare sung-lyrics *versus* spoken-lyrics in relation to sound distracters. Distracter sound manipulations for Study I included an instrumental version, a familiar sung-lyrics version, an unfamiliar sung-lyrics version, and a spoken-lyrics version of familiar nursery rhyme verses. This represented a departure from previous work in the context of short-term memory wherein simple tone sequences (Jones & Macken, 1993; Macken et al., 2003) or vocal and instrumental versions of two songs (Alley & Greene, 2008) have been presented. In addition, a large sample (294 participants) across a wide age-range (18-92) contrasted with previous studies (e.g., 60 students, Alley & Greene, 2008; 24 students, McCorkell, 2012; 24 students, Pring & Walker, 1994) allowed deliberation as to the generalizability of findings.

Findings that task-irrelevant sound generally impairs melody retrieval accuracy relative to quiet is consistent with a raft of previous findings using non-music tasks such as serial recall (Beaman & Jones, 1997; Colle & Welsh, 1977; Jones & Macken, 1993; Perham & Sykora, 2012; Salamé & Baddeley, 1982, 1989). Moreover, findings that sound does not need to be speech to produce disruption also coheres with previous work demonstrating disruption from instrumental music (Schlittmeier, 2008) as well as vocal music (Alley & Greene, 2008; Iwanaga & Ito, 2002; McCorkell, 2012). However, in contrast to previous studies reporting greater, or comparable, disruption of serial recall performance by speech than by non-speech sounds (Jones & Macken, 1993; Körner et al., 2017), in Study I the opposite is true: melody produced more disruption than spoken-lyrics.

In addition, that familiar instrumental melody was no more disruptive than unfamiliar melody is consistent with previous work on serial recall (Silverman, 2010). However, it is inconsistent with findings of Pring and Walker (1994) who reported nursery rhyme melodies to be more disruptive than unfamiliar classical music. Pring and Walker (1994) proposed that the additional disruption arose from automatic activation of lyrics associated with familiar nursery rhyme melodies as compared with unfamiliar classical music. However, they failed to demonstrate an activation of lyrics on an implicit memory test. Furthermore, the rather poor match of the familiar nursery rhyme melody and unfamiliar classical music on several musical criteria could alone explain their results (see Chapter II, section 2.6.3). Further, results from Study I show that a combination of familiar melody with unfamiliar or familiar lyrics disrupted accuracy of humming performance to a greater extent than unfamiliar lyrics and unfamiliar melody combined. Unfortunately, these results cannot be compared with any previous findings since, to the author's knowledge, no study has used analogous sound conditions.

Finally, a striking finding from the analysis of OST for melody production demonstrated that a combination of familiar melody with familiar lyrics slowed production to a greater extent than all other conditions. Again, the author is unaware of any analogous findings within the literature: response times are rarely measured for production of items or sequences within irrelevant sound paradigms involving production of multiple items from short-term memory. Theoretical interpretation of the reported findings forms the focus of the subsequent section.

3.5.1 Interference-by-content *versus* interference-by-process.

To recap, the focus of this current study was on structural *versus* dynamic views of the interference-forgetting relationship. The structural view proposes that interference arises because an encountered stimulus or event represented in memory is similar in some way to a target event, be those events verbal or non-verbal (for various examples see Baddeley, 1986; McGeogh, 1942; Mensink & Raajmakers, 1988; Salamé & Baddeley, 1982). Results of Study I seem altogether at odds with such an interference-by-content account.

A classic example of interference-by-content can be found within the WM model (Baddeley, 1986) wherein phonemes of irrelevant items from auditory origin gain direct access into the "phonological" store responsible for retaining those items over the short-term with the aid of the articulatory loop that recycles and revivifies decaying representations of those items. An early view was that a speech filter prevented non-speech items without phonological content from entering the phonological store and causing interference (Salamé & Baddeley,1982) and the degree of disruption of short-term memory for sequences of visual items was attributed to the degree of overlap between items of visual and auditory origin, within the phonological store (Salamé & Baddeley, 1982). Findings that non-speech sounds produce disruption of melody recall in the current study is inconsistent with previous account of the irrelevant sound effect wherein a selective filter only lets speech sounds within the phonological store (Salamé & Baddeley, 1982). However, it is consistent with a revised account within which auditory material that is sufficiently speech-like gains access to the phonological store, thereby assuming disruptive potency (Salamé & Baddeley, 1989). At odds with this interference-by-content account, however, is that non-speech melody produced greater disruption than speech alone. This suggests that the phonological (not lexical-semantic) content of the sound was responsible for disruption of melody retrieval from long-term memory. Similarly, that familiar melody in combination with lyrics (regardless of their familiarity) drove an additional disruption of melody retrieval is further at odds with the phonological store account.

To address whether the interference-by-content approach holds currency within the melody retrieval setting, several assumptions are required. One assumption is that there is an interplay between retrieval from long-term memory and temporary storage within the phonological loop of working memory, whereby retrieved lyrics are represented within the phonological store wherein they are held prior to production in the form of humming. If this view was correct, then the prediction would be that phonological similarity between target lyrics and those within the sound that are automatically represented within the phonological store, should determine disruption. However, there is no reason why familiar lyrics should be any more phonologically similar to a target melody than unfamiliar lyrics, yet the former were more disruptive (unless the lyrics were accompanied by a familiar melody).

Moving away from the WM model, the general construct of interference supposes that similar representations within the same memory space or store determines disruption. Assuming again that retrieval of melody from long-term memory involves some representation within a memory system prior to production, it is generally viewed that similarity between target and to-be-ignored representations should drive the disruption. However, similar problems arise for this perspective. For example, to explain why familiar melody accompanied by familiar or unfamiliar lyrics impairs melody production more than unfamiliar lyrics and unfamiliar melody combined, one would have to argue that similarity in representation of the former to a target melody is somehow greater than for the latter. It is difficult to envisage how this could be possible: great care was taken to match unfamiliar and familiar melodies and lyrics so at the objective level, between-sequence similarity (between target sequence and to-be-ignored sequence) was no different and yet familiar melody with accompanied lyrics was more disruptive than unfamiliar melody and unfamiliar lyrics combined. Thus, a structural, interference-by-content account of the interference-forgetting relationship in Study I appears unlikely.

A dynamic view of the interference-forgetting relationship is assumed by the process-oriented view (e.g., Marsh et al., 2009). In this view, processing necessitated by a focal task and that carried out on the irrelevant sound, rather than between-sequence similarity *per se.* determines disruption. Within the serial recall setting pre-attentive processing of changing elements in an acoustically varying sound, impedes deliberate vocal motor-planning process required for effective focal task performance (Hughes & Jones, 2005; Jones & Macken, 1993). The account readily explains the changing-state effect whereby changing-state sounds, as compared with steady-state sounds, produce greater disruption while the latter produce little, or no disruption, relative to quiet (Jones et al., 1992). The principle of interference-by-process in its original form, as applied to serial recall, would seem at odds with the pattern of data reported for Study I. First, speech distracters are often more disruptive than non-speech distracters, a finding that is typically attributed to greater acoustic variation in the former compared to the latter (Tremblay et al., 2000). However, contrary to this usual pattern, instrumental melody produced more disruption of melody retrieval than did speech in the current study. Second, although some musical excerpts are more acoustically complex than others (e.g., in terms of pitch, rhythm, dynamic) and have been shown to be more disruptive of serial recall (e.g., staccato produces more disruption than legato: Schlittmeier, 2008), acoustic variability of all musical items in this study (within Sound Type) was tightly controlled. Therefore, any additional disruption produced by familiar melody with lyrics (unfamiliar or familiar) over unfamiliar melody with unfamiliar lyrics cannot be attributable to greater acoustic variation and hence pronounced (acoustic) interference-by-process (cf. Jones & Tremblay, 2000).

An apparent collapse of the acoustic interference-by-process explanation to account for Study I results may be because processing used to retrieve melody is not a measure of seriation in its purest form. Like serial recall, melody production requires sequence processing: serial order of pitch forms a melodic contour which is of paramount importance in melody recall (Müllensiefen & Wiggins, 2010). Thus, fluctuation of acoustic properties of sound may well account for disruption of melody production by sound *per se*. This is since the necessitation of seriation as a focal task strategy renders it vulnerable to the changing-state effect. However, since a steady-state sound condition was not used as a comparison to melody and lyrics sequences it cannot be determined whether the disruption exerted by those sounds is due to their changing-state characteristics. In light of the foregoing some further specification of the parameters of interference-by-process is required.

3.5.2 The perceptual-gestural account.

The perceptual-gestural account (Hughes & Jones, 2005; Jones et al., 2006, 2007; Woodward et al., 2008) encompasses the interference-by-process mechanism and is aligned with the embodied cognition approach (Wilson, 2002) according to which general purpose processes and systems are co-opted for memory performance. In the context of serial recall, the skill of speaking is suitable for linking items that have no pre-existing ordered relationship (e.g., syntax/grammar) between them. This enables the embodiment of items in a temporally extended motor-plan. However, the skill of speech is an abstract entity that has underspecified action parameters. Initial un-population of the content of action parameters (movement of the vocal tract) means that any sequentially ordered irrelevant input threaten to populate, or indeed assume control of, the motor plan. Pre-attentive serial organization of acoustic elements of changing-state sound (as compared to steady-state sound yields a candidate for populating the motor-plan and thus disrupts serial recall (i.e., the changing-state effect; Jones & Macken, 1993). Importantly, the vocal-motor system is also used to represent melody due to the intricate relationship between perception of music and activation within the motor system (Leman, 2007; Sievers et al., 2013) that permits reproduction of melody. Furthermore, the vocal motor-system is involved in auditory imagery (Baddeley & Andrade, 2000; Reisberg et al., 1989) particularly for melody/song (Beaman et al., 2015; Smith et al., 1995). According to the perceptual-gestural view then (Hughes & Jones, 2005; Hughes et al., 2009), an intricate relationship exists between musical perception and motor-planning as it does between perceptual organization and motor-planning (in serial recall). Based upon this idea it was theorized that the vocal-motor system—responsible for planning and sequential production of melody—due to its under-specification of action parameters, could be disrupted by the presence of task-irrelevant sequences. From this platform, an outline of the impact of irrelevant sequences on melody production accuracy can be formulated.

First, it can be assumed that there is a basic form of disruption whereby any internal representation of a temporally-extended object (e.g., changing-state speech or changing-state nonspeech) as part of the perceptual streaming (or schema-driven) process, can interfere with melody planning and production process required by the humming task. This is because representation of the sound threatens to assume control of motor-programming of the target song. Second, a more specific form of disruption can be hypothesized: one that emanates from an earlier stage within the melody production process whereby a familiar song is selected among competitors within the musical lexicon for planning prior to production (Peretz & Zatorre, 2005; Peretz et al., 2009). This activation of competitors within the musical lexicon impairs selection of target melodies. A distracter comprising unfamiliar melody and unfamiliar lyrics does not activate a competitor (or may do so only weakly) within the musical lexicon because there is no pre-existing representation of that melody. However, distracters that comprise one or both elements that are familiar (e.g., unfamiliar melody with familiar lyrics; familiar melody with unfamiliar lyrics) produce more disruption than a distracter comprising unfamiliar melody and lyrics because they activate representation of song within the musical lexicon that acts as a competitor for the target melody. Activation of a competitor leads to a greater competition at the motor-planning stage for the target sequence. Presentation of a familiar melody with familiar lyrics produces faster recognition and activation of a competing melody within the lexicon that produces an initial impairment in time taken to recover the target melody (onset time) but henceforth produces as much disruption as familiar melody with unfamiliar lyrics and unfamiliar melody with familiar lyrics at the planning stage.

In the current study melody on its own was not enough to activate competition with the target melody (but see Chapter V-VI for a differing outcome). However, it could be argued that for Study I an unfamiliar matched melody, that contains the same rhythmic pattern as the original melody, unlike unfamiliar lyrics, is not a totally unfamiliar shape due to music's multifaceted structure. Although pitch formed a different melodic contour, harmonic progression, phrase structure, metre, and lilt were the same, and these elements, being fundamental to the original melody, would be stored in musical semantic memory, the musical lexicon (Peretz et al., 2009), awarding a degree of familiarity.

In order to perceive a musical line, the motor-system—that includes body movement such as gesture— works in partnership with musical perception (e.g., Leman, 2007). Nursery rhyme target and distracter melodies/lyrics employed for this study being frequently encoded with gesture (Thomas, 1930). When an irrelevant melody/song is perceived it has capacity to activate motor-system already required to produce similar neural responses to overt listening (e.g., King, 2006), this pathway could also operate for spoken-lyrics (with associated melody).

In the irrelevant melody only condition, lyrics usually associated with the target melody and, arguably, having been encoded together with melody when acquired, were missing but may have been activated by the melody. This supports findings noted by Peretz et al. (2004b), and Pring and Walker (1994). It was audibly obvious from live performance and audio recordings (although not formally assessed) that, on occasion, during the familiar distracter condition, there was considerable cueing of lyrics to trigger target melodies, participants using the inner voice (sub-vocal rehearsal) acting in partnership with the inner ear (store) to generate the target melody (Smith et al., 1995) sub-vocal control processes actively supporting mental rehearsal. In some cases, lyrics were used to produce the rhythmic pattern on a monotone without recall access to melodic contour. As all irrelevant familiar instrumental melodies contained associated lyrics this could, in part, have contributed to lower scores as compared to the spoken-lyric condition, contrary to McCorkell (2012) and Salamé and Baddeley (1989), or the ability of executive processes to override inhibition resulting from previously inhibited melodies containing similar features (Anderson et al., 2000).

The fact that familiar melody was not significantly more distracting than unfamiliar melody in Study I is puzzling. A possible explanation could be that some participants were unacquainted with a number of the nursery rhymes and so they were not recognized as familiar. Another possibility is that some participants' attention was captured by unfamiliarity of the pitch, albeit coupled with a familiar rhythm, that contributed to a search in the musical lexicon resulting in subsequent increased impairment.

3.5.3 Attentional capture.

An alternative to interference-by-process and interference-by-content accounts are attentional capture accounts (Bell, Dentale, Buchner, & Mayr, 2010; Cowan, 1995). These assume that auditory distraction is produced by an orienting response. In the context of serial recall, it is argued that changing-state sounds produce greater disruption than steady-state because each acoustic change within a sequence captures attention (Bell et al., 2010; Cowan, 1995). Within the context of serial recall this account cannot explain why tasks that do not require seriation are invulnerable to the changing-state effect (e.g., Beaman & Jones, 1997; Perham, Banbury, & Jones, 2007). Moreover, the account cannot adequately capture results reported here. Sequences comprising familiar lyrics and familiar melody, familiar lyrics with unfamiliar melody, and unfamiliar lyrics with unfamiliar melody produced more disruption of melody production accuracy than sequences of unfamiliar melody and unfamiliar lyrics. This effect cannot be attributed to a difference in the changing-state properties of the sequence since sequences were carefully matched, having the same vocal delivery, tonality, rhythmic and harmonic patterns, dynamic, with comparable undulating conjunct and disjunct melodic motion. Therefore, the only feature that was different across the four auditory sequences was distracter familiarity (melody, lyrics, or both). The unfolding contour of a familiar, compared to unfamiliar, melody/lyrics will be highly predictable resulting from a pre-existing neural model of the sounds (Koelsch, Vuust, & Friston, 2018; Pearce, 2018). Humans have the ability to consider a number of predictions about future auditory events, although these may be modulated by experience, or culture. Given that perception and action are intertwined (Monroy et al., 2019), perceptual and motor systems may interact, the sensory system able to anticipate the predicted sound (Tillman, Janata, Birk, & Bharaucha, 2008) and drive the motor system to deliver it. ERP experiments have shown motor performance of musicians guided by prediction of learned internal models (Bianco, Novembre, Keller, & Scharf, 2016; Sammler, Novembre, Kopelsch, & Keller, 2013). In this view, the prediction of a familiar contour occurring within irrelevant melody/lyrics should not be as attentionally captivating but may create competition within the motor systems, with production of a different familiar melody with consequent decrease in performance.

Another factor that governs whether attention is captured from a focal task by irrelevant auditory stimuli is personal significance or valence of those stimuli. It is important to note that such attentional capture is specific to the sound's content and is considered a qualitatively distinct form of auditory distraction to the changing-state effect by proponents of the interference-by-process/perceptual-gestural approach (e.g., the "duplex mechanism" account; Hughes, 2014). One example of attentional capture by personally significant stimuli is the own name effect whereby presentation of one's own name compared to a yoked control name, captures attention away from a focal task (Conway et al., 2001; Marsja, 2011; Röer et al., 2013). Further examples are the taboo word effect whereby inappropriate words produce more disruption of serial recall than neutral words (Röer et al., 2017), and the distracter valence effect whereby serial recall is disrupted more by emotionally valent than neutral words (Marsh et al., 2018). While it is unlikely that material used for irrelevant sound in this current study was taboo for participants, its familiarity could nevertheless have been personally significant, possibly eliciting an effective response (Marsja, 2011; Röer et al., 2013) attracting attention away from the melody production task. Crucially, in this view, disruption by irrelevant sound is not a joint function of the properties of the sound and focal task processes. Rather, the effect is task-process insensitive (Marsh et al., 2018).

Although distracter sound conditions in this current study did not contain a sudden deviant sound or word, known to disrupt task performance (Hughes et al., 2007), unexpected changes to the unfolding expectation of a known melody by addition of novel lyrics may have contributed to a deviation effect, as the ear was hearing different words/pitch to those that the memory was generating/imagining. One further possibility is that rather than sounds comprising familiar lyrics with unfamiliar melody and unfamiliar lyrics with familiar melody activating a competitor within the musical lexicon—thereby producing interference with motor-planning—the incongruency between melody and lyrics produced attentional capture.

3.6 Musicians *versus* non-musicians

Some prior research literature has shown participants with musical training out-perform non-musically trained participants in tasks of speech/verbal and tone/music comparison (Gröussard et al., 2010; Munza, Berti, & Pechman, 2002; Pechmann &

Mohr, 1992; Schön et al., 2004; Thompson & Yankeelov, 2012). It could, therefore, be hypothesized that musically trained/active participants would have an advantage in melody recall, preceding training equipping them with more highly tuned auditory features for the process and production of tonal pitch. Contrary to studies showing enhanced performance in music tasks for musically trained participants (e.g., Berti, Münzer, Schröger, & Pechmann, 2006; Munzer & Pechmann, 2000; Thompson & Yankeelov, 2012), Study I identified no benefit from musical training or participation in a musical activity. No significant effect or interaction was found. Therefore, no advantages were manifested.

3.7 Integration or independence?

Although domain specificity for linguistic and musical processing has been shown to be integrated in memory (Schön et al., 2010; Serafine et al., 1984), results from Study I would appear to support a degree of independence (Besson et al., 1998; Besson & Schön, 2001; Peretz et al., 1994). As lyrics and melody combined within song, familiar or unfamiliar, impaired melody retrieval to a greater degree than spoken-lyrics, familiar or unfamiliar, it suggests that song is stored differently to verbal information in WM (Berz, 1995; Besson & Schön, 2001; Halpern & Müllensiefen, 2008; Schulze et al., 2011). Moreover, spoken-lyrics, familiar or unfamiliar, without melody, were less disruptive than familiar melody and sung-lyrics, suggesting a degree of storage separation. However, interaction between semantic and musical memory, lyrics supporting melody retrieval, albeit an observation rather than an empirical finding, suggests an overlap of processing (Gröussard et al., 2010; Koelsch et al., 2002; Peynircioğlu et al., 1998). In addition, Study I result indicates humming does require access to the phonological lexicon, as sung-lyrics distracters produced greater disruption to hummed melody retrieval than irrelevant melody without lyrics. As humming a familiar melody, with associated lyrics, appeared to cue retrieval of associated lyrics, this suggests access was required to the phonological lexicon, although this deduction does not concur with previous research (cf. Peretz & Coltheart, 2003).

3.8 Concluding summary and subsequent research

In summary, the first study in this novel series focused on retrieval of one component of song, namely melody. Ability to retrieve and perform a familiar melody was disrupted by the presence of all irrelevant sound conditions deployed against quiet, thus indicating the mere presence of task-irrelevant sound disrupts retrieval of melody. What is more, results of this study showed there is a general form of auditory distraction in the context of melody retrieval that is independent of melody familiarity. Moreover, results demonstrated sung-lyrics more disruptive to melody recall performance than instrumental melody which, in turn, was more disruptive than spoken-lyrics. Additional specific effects of auditory distraction were also observed. For example, familiar sung-lyrics paired with familiar or unfamiliar melody impaired melody retrieval to a greater degree than unfamiliar sung-lyrics paired with unfamiliar melody. The patterns of results obtained for Study I appear at odds with structural views of the interference-by-forgetting relationship, while consistent with the dynamic view whereby task-irrelevant auditory material competes with (and thus disrupts) retrieval of a target melody to the extent that it competes for retrieval and impinges on vocal-motor planning process used to produce a melody. Formulated within the context of short-term memory, this account suggests there may be some overlap between mechanisms responsible for short-term rehearsal of verbal material and production of melody from long-term memory that merits further empirical study.

Another facet of this current study concerned integration or independence of melody and lyrics. Study I result suggests integration/independence, but this deserves further empirical investigation. By varying response task demands, while keeping auditory distracters constant, more weight may be added to the integration/independence perspective. If the same results obtained when a focal task required retrieval of known lyrics in the absence of melody, this would support the integration of melody and lyrics view. However, if melody retrieval were impaired to a greater degree than lyrics retrieval, this would suggest some degree of processing independence between music and lyrics within song.

CHAPTER IV

EMPIRICAL STUDY II: DISRUPTION OF SEMANTIC RETRIEVAL BY IRRELEVANT SOUND

Abstract

Song is a hybrid form derived from two separate entities, melody and lyrics, yet arguably represented in memory as an integrated unit. Whether melody and lyrics are stored as an integrated unit and accessed together (Pechmann & Mohr, 1992; Schön et al., 2010; Serafine et al., 1984, 1986), or stored independently and accessed separately (Bonnel et al., 2001; Peretz et al., 2009; Peretz & Zatorre, 2005), is a topic of active debate within the literature on music cognition (Besson & Schön, 2001; North & Hargreaves, 2008). Prior research comparing music and language processing has focused mainly on syntactic structure (e.g., Alluri et al., 2013; Besson & Schön, 2001; Fiveash & Pammer, 2014; Fiveash et al., 2018; Kunert et al., 2015; Miranda & Ullman, 2007; Patel, 2008; Peretz & Zatorre, 2005; Schön et al., 2010), and neuropsychological studies have not yet reached a consensus as to neural specificity, integration, or shared musical semantic memory processes (Alluri et al., 2013; Peretz & Zatorre, 2005; Platel et al., 2003). Study I (retrieval/performance of known song melody, via humming) showed familiarity, when embedded with sung-lyrics, to be the more potent distracter to recall. This suggests increased demands on the vocal-motor system resulting from the involuntary processing of sound sequences. Study II explores retrieval/performance of known song lyrics from long-term memory, via speaking, during irrelevant melody, a cappella song, or spoken-lyrics auditory distracters, familiar and unfamiliar. The rationale was that finding the same pattern of auditory distraction as obtained for Study I would add weight to the argument that melody and lyrics are integrated (e.g., Schön et al., 2010). However, if a different pattern of results arose for the current study then this would indicate some separation of melody and lyrics processing in relation to task demands (e.g., Besson et al., 1998; Bonnel et al., 2001). At odds with the integration view, the results demonstrated a different pattern of disruption for lyrics as compared with melody retrieval—spoken-lyrics being more disruptive of Study II than of Study I—thereby supporting the independence view. This conclusion was reinforced by a comparative analysis of the current study and Study I which showed that melody retrieval was impaired to a greater degree than spoken-lyrics retrieval and that spoken-lyrics against melody were more distracting of spoken-lyrics retrieval than melody retrieval (Study I).

4.1 Introduction

Language and music are high-level, hierarchical human skills and both domains share specific analogous characteristics, such as the need for rules, the “grammar” of sound governing how sentences and melodies are developed from a combination of basic units, such as words, and notes of pitch, universally termed syntax (Fiveash, McArthur, & Thompson, 2018; Thaut, 2009). Therefore, lexical knowledge within both domains for memorized representations is necessary (Miranda & Ullman, 2007). The idea that the language domain is specific to those cognitive functions relating to understanding, production, and use of language (Chomsky, 1965), has been disputed following studies that have identified similar brain responses when comparing aspects of language and music, such as syntax and harmonic violations (Kutas & Hillyard, 1983). Furthermore, the idea of music-specific neural networks, such as pitch, following identification of disorders that isolate musical abilities from language abilities within the cognitive system, adds to the debate (Peretz, 2003). Processing lyrics and melody has been shown to recruit distinct mechanisms (Ayotte et al., 2000; Hébert & Peretz, 2001) selective disruption of one song component, following brain damage, adds weight to the idea that recognition of lyrics and melody in song is facilitated by separate pathways (Peretz & Coltheart, 2003; Peretz et al., 2004).

The focus of the current chapter centres on this debate surrounding the integration/independent processing of melody and lyrics and the research builds on the foundations laid by Study I. Shaped in part by the modular model of music processing (Peretz & Coltheart, 2003, see 4.1.1 Figure 4), the present study contributes to the discussion of melody-lyrics independence or integration by examining pathways for the processing and production of lyrics. Study II continues the departure from more typically-used recognition tasks within music cognition research and closely replicates the procedure used in Study I so as to demand vocal production (in this case the spoken production of lyrics) and permit cross-experimental comparison and analysis as well as consideration of potential overlapping and non-overlapping processing components within the modular model (see Figure 4). To be clear, Study II requires participants to retrieve lyrics from long-term memory by spoken performance. The task of speaking lyrics, unlike melody production via humming, does not necessitate the production of melody (however, its retrieval according to the integration view will be retrieved automatically e.g., Serafine et al., 1984). Shifting the demands of the focal task also enables further assessment of the generality of the interference-by-process view (Jones

& Tremblay, 2000). Focussing the task demands on semantic processing may render the task vulnerable to disruption via the mere presence of meaning within the task-irrelevant sound: A semantic interference-by-process (e.g., Jones et al., 2012; Marsh et al., 2009). Generally, in this approach, one might expect spoken-lyrics to be more disruptive of retrieval than melody, the opposite pattern to that found in Study I, however, this must again be qualified by the extent to which familiar melody governs the automatic, implicit retrieval of associated lyrics.

4.1.1 Representation of melody and lyrics in song.

The relationship between words and music is most potent in song. On hearing and recognizing a familiar song melody, for example "Happy Birthday" the lyrics are presumed to be processed in parallel within the language processing system (Peretz et al., 2004). It appears that the auditory input—not a system of gatekeeping—dictates the pathway to the musical lexicon via musical modules (e.g., pitch, rhythm) or language modules (Peretz et al., 2009; Peretz & Zatorre, 2005), with all song information being directed to all modules (Coltheart, 2001). Therefore, if the focus were to produce (sing) the melody "Happy Birthday", the melody representation, from the musical lexicon, would be paired with associated lyrics from the phonological lexicon (Zatorre et al., 2007). Speculatively, this process may operate if "Happy Birthday" was heard without its associated lyrics, raising the question if it would be possible to produce (speak) the lyrics without accessing the melody?

Based on the idea of a dual store for song memory (Samson & Zatorre, 1991) there are two discrete pathways for output of music and lyrics. A singing route leads to a store wherein words are embedded in melody, and a language route leads to a store where words are represented alone. Although studies requiring production of lyrics and melody are scarce, some support for this idea has been generated from PET results with normal participants showing increased activity in the left hemisphere during speaking as compared to singing, and vice-versa in the right hemisphere (Jeffries, Fritz, & Braun, 2003). However, no support for the notion that singing promotes word intelligibility compared to speaking has been identified in a non-fluent aphasic patient (Hébert, Racette, Gagnon, & Peretz, 2003). Hébert et al. concluded that verbal production, by singing or speaking, is facilitated by the same (impaired) language output system, but this route is distinct from the (spared) melodic route. This idea was supported from a case study (Peretz et al., 2004) showing a patient with intact "la, la" melody production, but severe speech planning disorder for speaking and singing. The authors, therefore, concluded that lyrics can be distinguished from melody in singing and suggest that autonomy of music and language processing extends from recognition to production tasks. In this view, melody may not have to be accessed in order to produce spoken-lyrics. However, it is less clear if music and speech are fully integrated in song rather than just associated.

Working memory and requisite skills needed to enable sung-lyrics production, according to the vocal sensorimotor-loop model (Dalla Bella, Berkowska, & Sowiński, 2011; Tsang et al., 2011), may also be applied to spoken-lyrics production. To enable singing/speaking output, aspects of pitch/tempo/words are retrieved from memory and then sent to motor control areas. Accurate vocal performance of melody/lyrics then requires a process of auditory feedback to be maintained to check and correct output. This deliberate auditory to motor-planning procedure is one mechanism through which disruption can occur if concurrent involuntary motor-planning is generated (Dalla Bella et al., 2015; cf. Hughes & Jones, 2005; see Study I).

Pathways of information flow for familiar song spoken-lyrics output (with associated melody) during irrelevant familiar/unfamiliar melody/song/speech in this series of study may best be explained with reference to the modular model of music processing (Peretz & Coltheart, 2003, Figure 4). There have been limited studies relating to production of verbal/melody output (see Berkowska & Dalla Bella, 2009; Dalla-Bella et al., 2011) consequently it is still unclear how these pathways adjust when the output demand changes, such as when required to speak rather than sing or hum. So, the following exposition involves speculating as to the pathways generated for spoken-lyrics output performance.

For example, for Study I, based on the singing route of this architecture (Figure 4, red arrows) a familiar irrelevant melody input (e.g., “Ten green bottles”) would have a strong pathway through the musical lexicon, wherein its representation would be activated. It would then undergo some perceptual processing before proceeding to vocal plan formation where it would be in direct competition with vocal planning to produce a familiar target melody (e.g., “Jack and Jill”). However, an unfamiliar melody (e.g., “My dog Sam”) would have a weak pathway (based on matched rhythm and harmony) from the musical lexicon to the vocal plan/singing output since a representation is, by design, highly unlikely to be found within the musical lexicon. Consequently it should be less disruptive relative to familiar melody. As all familiar melodies/lyrics in this series have associated lyrics/melodies, each would also generate weak activation between the musical and phonological lexicons producing low-level competition on lyrics/melody retrieval.

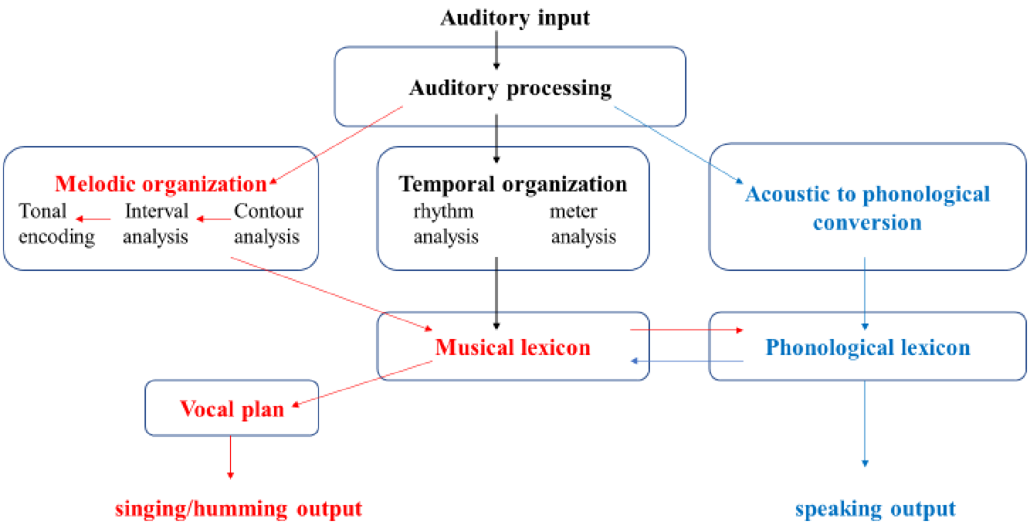


Figure 4. Cognitive neuropsychological model of processing pathway based on Peretz and Coltheart (2003). Note. ⟶ melody route ⟶ spoken-lyrics route.

Similarly, based on the speaking route (Figure 4, blue arrows), an irrelevant familiar spoken-lyrics input (e.g., "Ten green bottles") should have a strong pathway to the phonological lexicon, wherein its representation will be activated with conflict thereafter arising with vocal production to speak the target lyrics (e.g., "Jack and Jill"). However, the pathway for unfamiliar spoken-lyrics (e.g., "My dog Sam") to the phonological lexicon should be weak. In addition, there would be no activation support from the musical lexicon, and therefore less conflict generated.

When lyrics and melody are combined within song, both pathways should be activated, the strength dependent on familiarity of the lyrics/melody. For example, familiar lyrics with associated familiar melody should be more disruptive to the production of target lyrics than familiar melody alone, as it causes greater activation of competing lyrics within the phonological lexicon. However, unfamiliar lyrics may cause more disruption than unfamiliar melody since "words" are activated in the phonological lexicon (although these are not reciprocally activated by familiar melody, some activation may occur resultant from the matched musical components used in this series of study). Regarding the incongruity of familiar melody with unfamiliar lyrics, and vice-versa, familiar melody will activate the musical lexicon, (some associated lyrics activation also occurring in the phonological lexicon). Unfamiliar melody will cause weak activation of the musical lexicon, but familiar lyrics will be strongly activated in the phonological lexicon. Both unfamiliar melody and familiar lyrics should, therefore, impair retrieval of lyrics.

Regarding Study I (melody humming) the concept was that the motor-plan engendered by ignoring a sound could threaten control of the vocal apparatus resulting in poorer performance. This was attributed to the speech-planning mechanism required for melody production either being affected by the competition (interference from two concurrent motor-plans) and/or affected because the competing motor-plan actually assumed control of the vocal-motor planning and had to be removed (see Chapter III). This could also occur at the start of and during speech planning. When vocal output demands change, such as production of song lyrics as compared to production of song melody, but the presence of an irrelevant auditory condition is unchanged, based on this architecture (Figure 4) it may be possible to identify processing pathways more clearly. For example, humming requirements for Study I necessitated vocalization and motor-planning. An alternative mechanism to consider, however, is the sensorimotor model for singing output (Dalla Bella et al., 2015). In this approach, when an irrelevant melody simultaneously floods into the motor-plan during target motor-planning it has to be overridden. This can contribute to failure in monitoring the motor-output and impairment of the ability to adjust perceived contour according to self-monitoring feedback. Tentatively a familiar melody being intertwined with associated lyrics, could impair production of target lyrics in a similar manner. Although spoken retrieval of lyrics for this current study requires undertaking similar processes to melody retrieval, they are, arguably, lessened, as musical properties such as pitch contour and rhythm being irrelevant to spoken lyrics retrieval, may not be involved. Therefore, in this view, although production criteria differ, accuracy for spoken-lyrics production across irrelevant auditory conditions should be greater than for melody production.

4.1.2 Representation of music and language: Integration or independence?

The modular view of music processing outlined in 4.1.1 assumes some separability or independence of lyric and music processing. However, not all research agrees with the independence argument, many assume an integration of melody and lyric processing. A mass of research has accrued that compares music and language processing via fMRI and ERP imaging studies (e.g., Besson et al., 1998; Gröussard et al., 2010; Koelsch, 2005; Koelsch, Schulze, Alsop, & Schlaug, 2005; Miranda & Ullman, 2007; Patel, 2003, 2008; Serafine et al., 1984) and mainly focuses on structural aspects such as syntax/harmony. Here, a strong perception and processing relationship is evident between language and music whereby activation in Broca and Wernicke's areas indicate music harmonic processes are supported by language processes (Kunert et al., 2015; Patel, 2008). In contrast, the implication of distinct processing modules has arisen in neuropsychological studies showing dissociation for music and lyrics in people with brain damage (Hébert & Peretz, 2001; Peretz, 1996; Peretz et al., 1994; Risset, 1991; Samson & Peretz, 2005; Schlaug et al., 1995; for further detail see Chapter I).

Double dissociation for rule/memory in language and music (Miranda & Ullman, 2007) or in brain damaged patients such as amusia without aphasia and vice-versa (Peretz, 2003; Peretz, Champod, & Hyde, 2003) does not gel with the notion of a single processing system for music and language (Serafine et al., 1984, 1986), but points to a separation of processing modules (Peretz & Coltheart, 2003). Further adding to the debate, is the proposal that there exists a special store for songs following an amusic case wherein there occurred recognition of familiar song (presented on "la" without associated lyrics, but arguably holding embedded lyrics) but not recognition of familiar instrumental music (Steinke, Cuddy, & Jakobson, 2001). This finding also raises the question if there is a store for representations of instrumental music.

4.1.3 Impact of irrelevant sound on production of song elements.

Exploring the impact of different properties of task-irrelevant auditory distracters on production of melody (Study I) and lyrics (Study II) offers a window onto the nature of processing within the mnemonic systems responsible for the two properties of song (melody and lyrics), thereby shedding light on the integration *vs.* independence debate. Study I enquired as to whether the same principles underpinning distraction in short-term memory could apply to the specific situation of production of song melody from long-term memory in response to a song title cue. The results of that study were framed in terms of the interaction of motor-systems and dynamic processes that function to achieve task-oriented selective attention. Results from Study I showed all irrelevant sound distracters to impair melody recall, but more so when the sound contained sung-lyrics. Familiar sung-lyrics combined with unfamiliar melody was more disruptive to melody recall than unfamiliar melody combined with unfamiliar lyrics. Unfamiliar sung-lyrics combined with familiar melody was more disruptive than unfamiliar sung-lyrics combined with unfamiliar melody, although, a familiar sound was only more significantly distracting than an unfamiliar sound for the Unfamiliar Lyrics Sound Type Group. However, spoken-lyrics were not as distracting as melody irrespective of familiarity. Furthermore, a combination of familiar melody with familiar lyrics slowed onset time to begin production to a greater extent than all other conditions.

Those results were explained in terms of a conflict between vocal-motor processing and the automatic processing of the distracter sequences.

In keeping with the perceptual-gestural account (e.g., Jones et al., 2006, 2007), any changing-state sound sequence was held to interfere with melody production since this activity requires the serial organization of patterns of sound. More specific forms of disruption arise due to the familiarity of the irrelevant song based on its capacity to strongly activate a competitor song within the musical lexicon. Stronger activation within the musical lexicon feeds forward into the motor-planning stage for target production rendering the irrelevant sequence a more potent distracter thereby recruiting inhibitory mechanisms to minimise competition. The explanation that familiar song is a more potent distracter at the motor-planning stage because, as a familiar stimulus, it activates sequential ordering mechanisms more potently requires further empirical work[27]. According to the integration view, if melody and lyrics are integrated (Schön et al., 2010; Serafine et al., 1986), then the same pattern of disruptive effects attributable to task-irrelevant sound with different properties, observed in Study I, should be observed within Study II. Any deviation in the patterns of interference observed between Study I-II could, therefore, be taken as *prima facia* evidence for the independence view (e.g., Besson et al., 1998; Bonnel et al., 2001).

On the independence view, the modular model of music processing (Peretz & Coltheart, 2003) appears to predict differences between the disruptive impact of task-irrelevant sound as a function of the demands of the focal task (e.g., melody production via humming and lyric production for spoken output). A salient difference is that the vocal plan for singing and humming against speaking, requires melodic, pitch, spectral and temporal information (assuming participants do not covertly sing to decant lyrics into spoken output). If spoken production of lyrics to a song title does not entail activation of melody, then activation within the musical lexicon should give rise to access to the lyrics within the phonological lexicon that can then be produced. The susceptibility to distraction by task-irrelevant sound would be predicted to arise from sound streams that activate the musical lexicon/phonological lexicon. In this way, a key difference between Study I-II is that production of spoken lyrics (Study II) should be particularly prone to disruption from task-irrelevant lexical content, regardless of the presence of accompanying melody and familiarity, and less prone to disruption by

[27] This will be undertaken in Study III wherein it is investigated whether a familiar song component (e.g., melody) has a greater impact on the speech-planning mechanism *per se* as in the case of short-term serial recall.

melody. Note that this prediction pattern is the reverse to that observed for production of melody via humming (Study I) and thus dissociation should occur within a combined analysis.

Within the modular model of music processing (Peretz & Coltheart, 2003; Figure 4), it is possible to make more specific predictions regarding disruption effects observed from the task-irrelevant sound conditions. For example, in the case of familiar lyrics and familiar melody, disruption should occur within the phonological lexicon due to the task-irrelevant lyrics and the persistence of competitiveness of these lyrics could be promoted through the bidirectional links between the music and phonological lexicons. As a contrast, unfamiliar lyrics and unfamiliar melody should produce less activation within the phonological lexicon with most of this produced by unfamiliar lyrics. Unfamiliar lyrics coupled with a familiar melody might be expected to be no more disruptive than unfamiliar lyrics coupled with an unfamiliar melody. This is because the disruption produced to spoken-lyric retrieval in both cases should be attributable to the unfamiliar lyrics producing activation within the phonological lexicon. In this case, any activation via familiar melody of associated lyrics within the phonological lexicon in the unfamiliar lyrics with familiar melody sound type, might be inhibited by the activation of the unfamiliar lyrics. In the case of spoken distracters, the disruptive performance attributable to unfamiliar and familiar lyrics might be expected to be the same since in neither case is the activation of irrelevant lyrics within the phonological lexicon primed by the bidirectional connection between the music lexicon and the phonological lexicon through the presence of familiar melody. The latter case assumes an asymmetry in the bidirectional connection between music and phonological lexicons whereby spoken lyrics do not activate associated melody to the same extent as the reverse (cf. Asaridou & McQueen, 2013).

To summarize, Study II aims to examine language processing pathways through probed investigation of long-term memory for the lyrics of familiar song, via target lyrics speaking performance, in the presence of different irrelevant familiar and unfamiliar instrumental melody, sung-lyrics and spoken-lyrics. Building on Study I, the same irrelevant sound conditions will be used for Study II as a way of unpicking whether lyrics and associated melody, routinely encoded simultaneously as song (Peretz et al., 2004; Thomas, 1930) are stored and accessed dependently—integrated, as shown if an exact pattern of disruption occurred to that produced for Study I. A different pattern of results from Study I would suggest melody and lyrics were stored and accessed separately—independent within memory. This current study is the first to address the impact of concurrent, irrelevant, familiar and unfamiliar song distracters on the recall of complete verses of known lyrics from long-term memory.

4.2 Experiment 2: Method

4.2.1 Participants.

Two hundred and eighty-eight participants comprised of 54 men and 234 women aged 18-90 (*M* = 46.53, *SD* = 22.41[28]) undertook Experiment 2. Participants were recruited from U3A groups, Soroptimist, Rotary, choirs, universities, colleges, nurseries, community groups, and personal contact. (In total, 296 participants were tested, however, owing to a lack of knowledge of the target nursery rhymes [zero known melodies registered for any total trial condition] eight were eliminated from the study and not replaced). Ethical approval (Appendix A) and reimbursement arrangements detailed in Chapter III.

Cognition.

In order to continue to control for age-related cognitive decline, adults aged over 50 years were asked to complete the Addenbrooke's Cognitive Examination Revised (ACE-R) which also gives mini-mental state examination (MMSE) information[29]. The scoring ranged between 82-99 (*M* = 93.39, *SD* = 4.19) with MMSE scores between 25-30 (*M* = 28.90, *SD* = 1.11) falling within the range of expectation on the measure, which indicates positive cognitive functioning. No participants scored <82 which is the suggested cut-off point for dementia sensitivity. A repeated measure mixed analysis of variance (ANOVA) measuring mean recall scores across the three Sound Conditions against participants' cognitive scores showed that the effect between accuracy of lyrics recall and cognition[30] was non-significant, although it lay on the cusp, $F(2, 284) = 2.99$, $MSE = .140$, $p = .052$, $\eta_p^2 = .021$. Moreover, the Sound Condition × Cognition interaction was also non-significant, $F(2, 284) = 1.343$, $MSE = .140$, $p = .263$, $\eta_p^2 = .009$. It would seem therefore, that levels of cognitive ability did not detrimentally influence semantic recall in the presence of sound distracters (Appendix D). In keeping with Study I, a visual acuity test (task font 22-24) confirmed acceptable vision and a test for hearing level, conducted via an audiometer, recorded hearing levels > 30dB at 0.5-4.0 kHz for all participants indicated normal hearing[31].

---

[28] The mean falling within the same mode range as Experiment 1 (Age 45-49).
[29] Format and procedure detailed in Chapter III.
[30] From ACE-R data.
[31] Normal hearing established at 0-25 dB, some mild hearing loss at whisper level recorded at 20-40 dB.

Musical culture.

All participants completed a demographic response sheet. Responses showed 91% of participants indicated a "Western" musical culture, with 6.9% indicating joint "Asian" and 2.1% joint "African" musical cultures. This slight rise in a non-western cultural background reflected increased participation for Study II from the culturally diverse UCLan[32] community. The eight participants eliminated from the study (having scored zero known for one trial) indicated a non-western musical culture.

Musical training and participation.

An ANOVA comparing speaking scores of all participants with prior musical training, as identified on their demographic response sheet, indicated a non-significant effect between Sound Condition and prior musical training, and between Sound Condition and musical participation. However, there was a significant between-participants effect for musical training, $F(1, 284) = 7.707$, $MSE = .962$, $p = .006$, $\eta_p^2 =.026$, and choir participation, $F(1, 284) = 8.890$, $MSE = .962$, $p = .003$, $\eta_p^2 = .030$, possibly suggesting a deeper encoding and retention, nursey rhymes being a popular medium for early musical training. However, current instrumental participation was non-significant, $F(1,284) = 1.455$, $MSE = .962$, $p = .229$, $\eta_p^2 =.005$.

4.2.2 Design.

Experiment 2 was conducted as a mixed 3 (Sound Condition) × 4 (Sound Type) design, the dependent variable being accuracy of song lyrics recall through speaking. A further mixed ANOVA was conducted to establish prior knowledge of the target and distracter lyrics from a Recognition test, and an additional 3 (Sound Condition) × 4 (Sound Type) mixed ANOVA examined onset time taken to begin production of the target lyrics. The within-participants variable of Sound Condition was classified into three levels, quiet (no distracter), familiar distracter present, and unfamiliar distracter present. The between-participants variable of Sound Type was classified into four levels, Familiar Lyrics, Melody, Speech (spoken-lyrics), and Unfamiliar Lyrics, and participants were allocated alternately to each Sound Type (Group) see Table 4.1. The order of presentation of distracter conditions followed the same design as Study I.

[32] UCLan (University of Central Lancashire).

Table 4.1

Combination of experimental Sound Conditions for Sound Type Groups

Sound Type Group	*n*	distracter variable
All	288	quiet
Familiar Lyrics	72	familiar sung-lyrics (familiar melody)
		familiar sung-lyrics (unfamiliar matched melody)
Melody	72	familiar melody (no lyrics)
		unfamiliar matched melody (no lyrics)
Speech	72	familiar spoken-lyrics (no melody)
		unfamiliar matched spoken-lyrics (no melody)
Unfamiliar Lyrics	72	unfamiliar sung-lyrics (familiar melody)
		unfamiliar sung-lyrics (unfamiliar matched melody)

4.2.3 Materials and apparatus.

All forty western culture target and distracter nursery rhymes, traditional, and patriotic songs, were identical to those used in Study I (melody humming). They were assigned to the same four Sound Type Groups and were delivered through the E-Prime 2.0 Psychology software tool (for details of stimuli specifics see Chapter III, 3.2.3). To reiterate, however, unfamiliar matched song lyrics comprised a syntactical poem. Rather than unfamiliar lyrics following a retrograde familiar word pattern (Hébert, Peretz, & Gagnon, 1995; Peretz et al, 2009) novel words were created to mirror the conspicuous repetitive word patterns epitomized in nursery rhymes. These novel lyrics were specifically created to also follow the rhythmic pattern of the original rhyme. They comprised a predominantly syllabic setting with limited melisma, similar repetitive metre, and general syntactic structure of the original lyrics (see Chapter III 3.2.3 and example in Figure 5 below).

Oh where, oh where has my little dog gone?

Original nursery rhyme lyrics:

Oh / where, oh where has my / little dog gone, oh / where oh where can he / be? with his / ears cut short and his / tail cut long, oh / where oh where can he / be?

Matched lyrics:

I / saw a beautiful / red butterfly, ‘twas / sitting looking at / me. It just / raised its wings as it / said goodbye and / sailed away in the / sky.

Figure 5. Example of original and matched rhyme lyrics showing 6/8 time bar divisions.

For all sung and spoken excerpts a female contralto voice, a trained musician, produced a live a cappella performance into a laptop computer via a line-in USB microphone, using the sound recording and editing-program Audacity. Speech was delivered on a monotone with limited vocal inflection and was paced to the tempo of the sung version in keeping with Study I and previous research (Simmons-Stern, 2012). In this way sung and spoken tracks were of equal duration.

4.2.4 Procedure.

To assess whether melody and lyrics of song are integrated or represented independently various combinations of melody and lyrics were manipulated and delivered, to predict if retrieval was limited or affected through impairment due to the presence of a distracter. Here, the focal task requirement was to speak the lyrics of a well-known rhyme in response to a visually presented title (e.g., “Jack and Jill”) whilst ignoring another well-known melody, sung- or spoken-lyrics (e.g. "Ten green bottles”) or an unfamiliar/novel instrumental melody, sung- or spoken-lyrics. All testing was undertaken individually in a quiet room in the presence of the researcher. The laptop computer internal microphone with a 20dB boost was used for recording and all participants wore a set of Sennheiser HD201 headphones attached to the computer to receive the irrelevant audio. Irrelevant sounds averaged 62dB (A). To ensure standardisation of instructions to participants instructions to “Press the spacebar when you are ready to progress” accompanied by an audible bleep were consistent for each movement through the program.

Following two practice trials without distracters, all participants experienced two distracter sound condition blocks (familiar and unfamiliar), and a quiet (no-sound) block. The order was counterbalanced to reduce demand characteristics, and song-titles within each block randomly determined. This enabled 72 participants in the Familiar Lyrics Group to experience familiar melody (familiar sung-lyrics) and unfamiliar matched melody (familiar sung-lyrics), 72 participants in the Melody Group to experience familiar melody (no lyrics) and unfamiliar matched melody (no lyrics), 72 in the Speech Group experienced familiar spoken-lyrics and unfamiliar spoken-lyrics, with 72 in the Unfamiliar Lyrics Group having familiar melody (unfamiliar sung-lyrics) and unfamiliar matched melody (unfamiliar sung-lyrics). For example, a participant in the Melody Group was required to speak the lyrics of ten target songs in each of three condition blocks, according to the title visible on the computer screen, while accompanied by to-be-ignored sound comprising either a familiar melody, an unfamiliar melody, or in quiet. Target titles were taken from song-set A or B or C depending on the program melody set order in the six orders of distraction. Both familiar and unfamiliar matched distracter melodies were taken from melody set D (see Appendix C). Participants in the Familiar Lyrics Group, Speech Group, and Unfamiliar Lyrics Group followed the same procedure with their respectively different distracters. The visual appearance of the target song title and the distracter sound were set to begin simultaneously. Participants' speaking performances were recorded directly into the computer via its inbuilt computer microphone. Response times to begin production of target lyrics were also recorded by the E-prime program. Finally, after participants had undertaken all experimental trials, a Recognition test sought to determine previous knowledge of target and distracter lyrics. The titles of all 40 songs (including 10 distracter songs) were shown individually via the computer screen. Participants were instructed to indicate whether the rhyme lyrics were familiar or unfamiliar. They were instructed to press the "Y" button on the computer keyboard for Yes if familiar, and the "N" button for No if not familiar to them. If the "Y" button was pressed they were then instructed to speak the lyrics of the target song again. If they pressed "N" the next rhyme title immediately appeared. All speaking attempts from each experimental trial was recorded through the computer and converted by the E-prime program into individual .WAV files.

4.2.5 Assessment.

For Study II only recall accuracy of nursery rhyme lyrics was assessed. Speaking attempts from each of the experimental trials were recorded through the computer and converted by the E-prime program into individual .WAV files. All target lyrics performances were measured using the same scale as for Study I (Chapter III) but with reference to the number of correct lyrics in each bar as opposed to correct notes of pitch, which ensured a consistent tool of measure across both experiments. Performances were assessed using a scale range from zero, inaccurate, to four, accurate and fluent, based on established musical performance assessment criteria. The scale was designed to allow a zero score for lyrics attempted but comprised of no rewardable material. Lyrics of the rhyme poems were divided into bars matching their sung version, the total number of bars for each song then divided into four sections of equal number to match criterion descriptors, see Table 4.2. These criteria provided validity for measure of exact purpose, word accuracy, without reference to other musical elements.

Raw data from each .WAV file of spoken-lyrics from each participant was listened to and assessed by the author/researcher according to criteria identified in Table

4.2. The .WAV files were also burnt to CD for two independent raters, trained teachers, to assess. They compared performances from half of each Sound Type Group for accuracy to the target lyrics.

Table 4.2

Assessment criteria for the measurement of speaking retrieval accuracy

Score	0	1	2	3	4
Criteria	inaccurate, no rewardable content	limited accuracy, repeated errors	broadly accurate, some errors	mainly accurate	accurate + fluent
total bars	bars correct	bars correct	bars correct	bars correct	bars correct
8	0	1-2	3-4	5-6	7-8
12	0	1-3	4-6	7-9	10-12
14	0	1-3.5	3.5-7	7-10.5	10.5-14
16-15	0	1-4	5-8	9-12	13-16
20	0	1-5	6-10	11-15	16-20

An intraclass correlation coefficient (ICC) showed an extremely high consistency between raters, Cronbach's $\alpha = .994$. Unlike Study I, (melody recall) it was not necessary to replicate the rating with musical and non-musically trained raters, as production of lyrics by speaking, as opposed to production of a melody by humming, was considered less prone to subjective assessment. Assessment from the audio recordings did not identify any outliers. Onset time, in seconds, measuring time taken to read the title, search retrieval of target rhyme lyrics from long-term memory, and begin production of target lyrics by speaking was also recorded.

4.3 Results

The main purpose of this empirical study was to identify the effect on accuracy of lyrics retrieval from long-term memory, and performance by speaking, from 288 participants in quiet conditions and in the presence or absence of a familiar or unfamiliar sound distracter, by comparing means for each of the three levels of Sound Condition.

Lyric retrieval via speaking performance was affected by the presence of both familiar and unfamiliar task-irrelevant sound. Accuracy was higher for all Sound Type Groups in the quiet condition (although slightly lower for the Unfamiliar Lyrics Group). However, across all Sound Types, semantic recall accuracy was impaired to a greater extent in the presence of a familiar distracter as compared to an unfamiliar distracter, (see Figure 6). These results suggest the Familiar Lyrics Group to be the lowest scoring across Sound Types (although not in quiet) irrespective of melody familiarity.

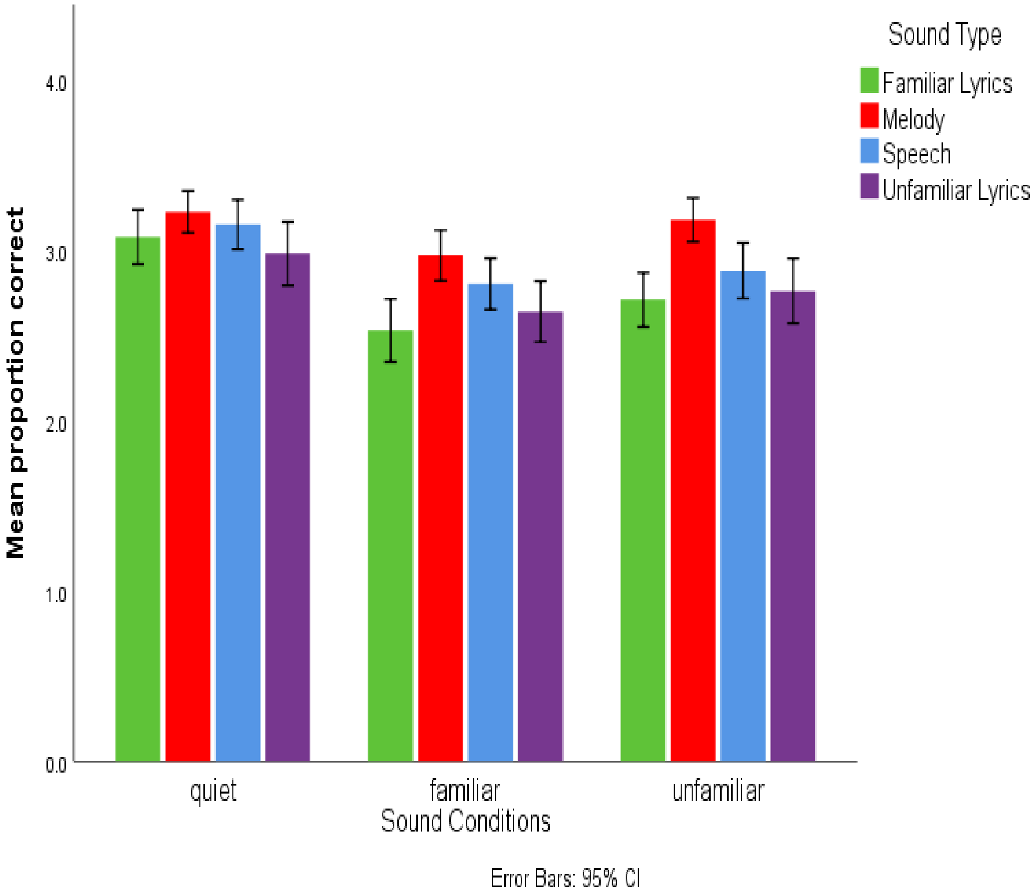


Figure 6. Mean proportion retrieval accuracy for lyrics speaking across the quiet, familiar, and unfamiliar Sound Conditions according to the Sound Type Group variable.

Results from a 3 (Sound Condition: quiet, familiar, unfamiliar) × 4 (Sound Type: Familiar Lyrics, Melody, Speech [spoken-lyrics], Unfamiliar Lyrics) mixed ANOVA showed a significant main effect of Sound Condition, $F(2, 568) = 54.377$, $MSE = .187$, $p < .001$, $\eta_p^2 = .161$, indicating lyrics retrieval was affected by the presence of familiar and unfamiliar task irrelevant sound: sung, spoken, or instrumental. There was also a between-participants main effect of Sound Type, $F(3, 284) = 5.730$, $MSE = 1.036$, $p = .001$, $\eta_p^2 = .055$, indicating significant differences in recall accuracy between the four different Sound Type Groups. Pairwise comparisons were not computed as the pattern of performance within groups and across groups differed with the Sound Conditions and Sound Type factors, respectively, as revealed by a significant Sound Condition × Sound Type interaction, $F(6, 568) = 2.371$, $MSE = .187$, $p = .029$, $\eta_p^2 = .024$. This indicates that the Sound Type was influential in determining the accuracy of task performance, the pattern of performance being dependent on both Sound Type (the nature of the distracter material; melody *vs.* sung-lyrics *vs.* spoken-lyrics) and Sound Condition (quiet, familiar, unfamiliar).

Decomposition of the interaction via simple effects analysis[33] was undertaken to better understand how performance was disrupted as a function of Sound Type and Sound Condition. The first phase of analysis compared means for Sound Conditions (quiet, familiar, unfamiliar) as a function of Sound Type (Familiar Lyrics, Melody, Speech, Unfamiliar Lyrics). Initial comparisons were focused on determining whether the presence of sound, regardless of its familiarity, determined disruption by comparing quiet and unfamiliar sound conditions within each Sound Type. Performance was poorer in the unfamiliar condition compared to the quiet condition for Familiar Lyrics (MD = .368, SE = .073, p < .001; 95% CI [.223, .513]), Speech (MD = .272, SE = .073, p < .001; 95% CI [.128, .417]), and Unfamiliar Lyrics (MD = .219, SE = .073, p = .003; 95% CI [.075, .364]) Groups. This demonstrated that speaking accuracy performance was poorer when accompanied by sound containing lyrics, sung or spoken, and regardless of whether the sound comprised a familiar component (e.g., familiar lyrics) or not (e.g., unfamiliar lyrics). However, the difference between the unfamiliar condition and quiet was non-significant for the Melody Group (MD = .046, SE = .073, p = .533; 95% CI [-.099, .190]). This result indicates that speaking performance was not affected by the mere presence of an unfamiliar melody relative to quiet in the absence of lyrics.

Secondary comparisons investigated whether familiarity effects occurred within each group by comparing familiar and unfamiliar sound conditions. Performance was significantly poorer in the familiar compared with unfamiliar sound conditions for the Familiar Lyrics Group (MD = -.179, SE = .069, p = .010; 95% CI [-.315, -.043]), and for the Melody Group (MD = -.210, SE = .069, p = .003; 95% CI [-.346, -.074]), thereby demonstrating a familiarity effect. See Table 4.3 for all familiar *versus* unfamiliar pairwise comparisons. (*Note*. 95% confidence intervals.)

---

[33] LSD reported as it maintains statistical power. See Chapter III, 3.2.5.

Table 4.3

Sound Condition comparisons: familiar melody (fam) versus (vs.) unfamiliar melody (unf) within Sound Type Group

Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Familiar Lyrics	fam *vs.* unf	-.179	.069	.010	-.315	-.043
Melody	fam *vs.* unf	-.210	.069	.003	-.346	-.074
Speech	fam *vs.* unf	-.078	.069	.260	-.214	.058
Unfamiliar Lyrics	fam *vs.* unf	-.121	.069	.081	-.257	.015

No such familiarity effect emerged for the Speech Group (p = .260), or the Unfamiliar Lyrics Group (p = .081). Thus, the production of target lyrics (e.g., 'Mary, Mary, quite contrary') was no more impaired by exposure to unfamiliar spoken-lyrics (e.g., 'Hold my hand'; Figure 1h[34]) as compared with familiar spoken-lyrics (e.g., 'London Bridge'; Figure 1g). Furthermore, the presence of unfamiliar lyrics (e.g., 'Hold my hand') accompanied by familiar melody (e.g., 'London Bridge'; Figure 1e) produces no more disruption to the production of a target lyric (e.g., 'Mary, Mary quite contrary') than unfamiliar lyrics accompanied by unfamiliar melody (e.g., 'Hold my hand'; Figure 1f). This suggests that familiarity provided by melody—within a to-be-ignored stimulus—produces disruption relative to when melody familiarity is not present or heard in combination with un-associated lyrics.

The second phase of analysis explored differences across Sound Type Group as a function of Sound Condition. There was a significant difference during the baseline (quiet) condition between the Melody Group and Unfamiliar Lyrics Group (*MD* = -.244, *SE* = .110, p = .027; 95% CI [.028, .461]), confirming recall performance to be superior for the Melody Group as shown in Table 4.4.

[34] Distracter examples Figures 1a-h shown in Chapter III 3.2.3.

Table 4.4

Sound Condition comparison across Sound Type Group

Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Melody *vs.* Unfamiliar Lyrics	quiet	.244	.110	.027*	-.028	.461
Familiar Lyrics *vs.* Melody	familiar	-.439	.118	< .001*	-.670	-.207
Familiar Lyrics *vs.* Speech		-.272	.118	.021*	-.504	-.041
Melody *vs.* Speech		.167	.118	.158	-.065	.398
Speech *vs.* Unfamiliar Lyrics		.163	.118	.168	-.069	.394
Unfamiliar Lyrics *vs.* Familiar Lyrics		.110	.118	.352	-.122	.341
Melody *vs.* Unfamiliar Lyrics		.329	.118	.005*	.098	.561
Familiar Lyrics *vs.* Melody	unfamiliar	-.469	.115	< .001*	-.696	-.243
Familiar Lyrics *vs.* Speech		-.171	.115	.139	-.397	.056
Familiar Lyrics *vs.* Unfamiliar Lyrics		-.051	.115	.655	-.278	.175
Melody *vs.* Speech		.299	.115	.010*	.072	.525
Speech *vs.* Unfamiliar Lyrics		.119	.155	.300	-.107	.346
Unfamiliar Lyrics *vs.* Melody		-.418	.115	< .001*	-.644	-.192

Note. * denotes significance.

In the context of unfamiliar sound, lyric production via speaking performance was superior when accompanied by melody as compared with speech (*MD* = .299, *SE* = .115, *p* = .010; 95% CI [.072, .525]). Thus, accuracy of speaking performance was superior when the target lyrics (e.g., ‘Mary, Mary, quite contrary’) was attempted in the presence of unfamiliar melody (e.g., ‘Hold my hand’; Figure 1b) compared to spoken-lyrics (e.g., ‘Hold my hand; Figure 1h).

Performance was poorer when unfamiliar melody was combined with familiar lyrics compared to unfamiliar melody alone (*MD* = -.469, *SE* = .115, *p* < .001; 95% CI [-.696, -.243]). Thus, lyrics speaking was poorer when the target lyrics (e.g., ‘Mary,

Mary, quite contrary') was attempted in the presence of an unfamiliar melody (e.g., 'Hold my hand') combined with familiar lyrics (e.g., 'London Bridge'; Figure 1d) compared to unfamiliar melody alone (e.g., 'Hold my hand'; Figure 1b). However, when an unfamiliar melody was combined with unfamiliar lyrics (e.g., 'Hold my hand'; Figure 1f) speaking performance was poorer compared to during an unfamiliar melody (e.g., 'Hold my hand'; Figure 1b) without lyrics (*MD* = -.418, *SE* = .115, $p < .001$; 95% CI [-.644, -.192]).

Within the context of familiar sound, the combination of familiar melody and familiar lyrics produced greater disruption to speaking recall accuracy than familiar melody alone (*MD* = -.439, *SE* = .118, $p < .001$; 95% CI [-.670, -.207]). Furthermore, combinations of familiar melody with familiar sung-lyrics produced more disruption than familiar spoken-lyrics (*MD* = -.272, *SE* = .118, $p = .021$; 95% CI [-.942, -.372]. Therefore, production of target lyrics (e.g., 'Mary, Mary, quite contrary') was impeded more by the combination of familiar melody and familiar lyrics (e.g., 'London Bridge'; Figure 1c) than by a familiar melody or spoken-lyrics (e.g., 'London Bridge'; Figure 1a and 1g respectively). In addition, combination of familiar melody and unfamiliar sung-lyrics produced greater disruption to lyrics speaking recall accuracy than familiar melody alone (*MD* = -.329, *SE* = .118, $p = .005$; 95% CI [-.561, -.089]). Therefore, speaking of familiar rhyme lyrics (e.g., 'Mary, Mary, quite contrary') was impeded more by the incongruity of unfamiliar lyrics (e.g., 'Hold my hand') with a familiar melody (e.g., 'London Bridge'; Figure 1f) than by a familiar melody alone (e.g., 'London Bridge'; Figure 1a).

Combination of familiar melody with familiar lyrics was no more disruptive than the combination of familiar melody with unfamiliar lyrics ($p = .352$). However, recall accuracy during unfamiliar sung-lyrics accompanied by familiar melody was lower than performance in the familiar melody alone condition ($p = .005$). Thus, target lyrics production (e.g., Mary, Mary, quite contrary) was more impaired by concurrently ignoring a combination of unfamiliar lyrics (e.g., 'Hold my hand') with familiar melody (e.g., 'London Bridge'; Figure 1e) than from ignoring a familiar melody alone (e.g., 'London Bridge'; Figure 1a).

To summarize, foregoing results showed that retrieval of target lyrics—as indexed by accuracy of speaking performance—was adversely affected by all Sound Conditions as compared to quiet. Although there was a tendency for familiar sound to be more disruptive than unfamiliar sound in the Speech and Unfamiliar Lyrics Group, a significant difference only arose for the Melody Group and the Familiar Lyrics Group.

Thus, a familiar melody was more potent at impeding speaking performance than an unfamiliar melody, and, crucially, familiar lyrics combined with familiar melody was more disruptive than familiar lyrics combined with unfamiliar melody. This suggests melody familiarity overrode lyrics familiarity if there was an incongruity (familiar lyrics but with unfamiliar melody). Absence of a familiarity effect (although close, $p = .081$) within the Unfamiliar Lyrics Group, however, may be as a consequence of familiarity with the lyrics dimension alone required to disrupt semantic retrieval. Melody was generally less disruptive than spoken-lyrics (Speech) but only significant with the unfamiliar Sound Condition.

While the combination of unfamiliar melody and familiar lyrics was clearly more disruptive than unfamiliar melody, and the combination of unfamiliar melody with unfamiliar lyrics was clearly more disruptive than unfamiliar melody, differences in relation to pairings within the context of familiar sounds were more diverse. While the combination of familiar melody and lyrics impaired recall accuracy more than familiar melody alone, and familiar spoken-lyrics, it was not more disruptive than familiar melody combined with unfamiliar lyrics, and this condition, although more disruptive than melody, was not more disruptive than familiar speech.

4.3.1 Recognition test.

Following distracter trials, and in accordance with the procedure used for Study I, a Recognition test was conducted in quiet conditions. All 40 nursery rhyme titles (including distracter rhymes) were presented again for lyrics recall controlling for prior knowledge. If the “Y” key on the computer keyboard was pressed by participants to indicate that the lyrics of the target song was ‘familiar’ to them, participants were asked to speak the words again. The total mean of “familiar” song lyrics acknowledged was 27.61 (*SD* = 9.4) indicating 69% knowledge of target and distracter lyrics. However, the number of known song lyrics attempted and recognized, from a possible 10 for each experimental trial, was 8.03 (*SD* = 2.47) in quiet, 7.99 (*SD* = 2.41) in the familiar sound condition, and 8.01 (*SD* = 2.41) in the unfamiliar sound condition, generating an 80% total mean of recognized nursery rhyme lyrics. This inconsistency may, in part, be attributed to some participant’s reticence in the Recognition test to commit to a “Y” if the complete rhyme was not known, whereas in trials partial retrieval of lyrics had been recorded before realisation that all the lyrics were either not known or could not be remembered. This discrepancy resulted in a slightly lower total percentage in the Recognition test than trials would indicate. However, 6.51 (*SD* = 2.28) from a possible 10 distracter poems were indicated as familiar.

A mixed ANOVA was conducted between the mean recall scores from the two sung-lyrics group distracter trials, the lowest scoring, against the mean recall scores from the final Recognition test in the absence of a distracter, with Sound Type Group as the between-participants factor (see Table 4.5), and this showed increased accuracy against quiet.

Table 4.5

Mean retrieval scores for familiar and unfamiliar Sound Conditions, with and without distracter, for Familiar Lyrics and Unfamiliar Lyrics Sound Type Groups

Sound Type Group	familiar distracter	Recognition without distracter	unfamiliar distracter	Recognition without distracter
Familiar Lyrics	2.53 *(.78)*	3.44 *(.48)*	2.71 *(.68)*	3.38 *(.67)*
Unfamiliar Lyrics	2.64 *(.76)*	3.30 *(.60)*	2.76 *(.81)*	3.23 *(.74)*

Results[35] showed a significant main effect of Sound Condition, $F(2.673, 758.995) = 182.989$, $MSE = .210$, $p < .001$, $\eta_p^2 = .392$, suggesting increased accuracy due to absence of distracter rather than lack of prior knowledge. There was also a significant between-participants effect of Sound Type, $F(3, 284) = 5.208$, $MSE = .1.086$, $p = .002$, $\eta_p^2 = .052$, indicating some difference in accuracy for the Sound Type Groups. A simple effects analysis revealed this to be between the Melody Group and all lyrics groups, see Table 4.6.

Table 4.6

Pairwise comparison for Sound Type in Recognition Test

Sound Type Group	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Familiar Lyrics *vs.* Melody	-.281	.087	.001	-.451	-.110
Speech *vs.* Melody	-.173	.087	.047	-.344	-.002
Unfamiliar Lyrics *vs.* Melody	-.310	.087	< .001	-.481	-.139

Pairwise comparisons showed significance (p's $< .010$) between all Sound Conditions except for between the two Recognition test conditions, ($MD = .030$, $SE = .039$, $p = .451$; 95% CI [-.047, .106]). This suggests increased scoring in the quiet conditions of the Recognition test from the two lowest scoring lyrics groups, compared to scoring in the distracter trials, was attributable to the quiet conditions of the Recognition test and not from lack of knowledge of target rhymes. In addition, there

[35] Huynh-Feldt reported following sphericity violation $\varepsilon = .891$

was a significant two-way Sound Condition × Sound Type interaction, $F(8.018, 758.995) = 3.154$, $MSE = .210$, $p = .002$, $\eta_p^2 = .032$. A simple effects analysis for Sound Condition as a function of Sound Type identified this effect to be driven by all lyrics Groups, but particularly by sung-lyrics Groups, see Table 4.7.

Table 4.7

Pairwise comparisons for Sound Condition across Sound Type between familiar and unfamiliar sound condition trial scores and Recognition (Recog) test scores

Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Familiar Lyrics *vs*. Melody	familiar	-.439	.118	< .001*	-.670	-.207
Familiar Lyrics *vs*. Speech		-.272	.118	.021*	-.504	-.041
Melody *vs*. Speech		.167	.118	.158	-.065	.398
Speech *vs*. Unfamiliar Lyrics		.163	.118	.168	-.069	.394
Unfamiliar Lyrics *vs*. Melody		-.329	.118	.005*	-.561	-.098
Familiar Lyrics *vs*. Melody	unfamiliar	-.469	.115	< .001*	-.696	-.243
Familiar Lyrics *vs*. Speech		-.171	.115	.139	-.397	.056
Familiar Lyrics *vs*. Unfamiliar Lyrics		-.051	.115	.655	-.278	.175
Melody *vs*. Speech		.299	.115	.010*	.072	.525
Unfamiliar Lyrics *vs*. Melody		-.418	.115	< .001*	-.644	-.192
Familiar Lyrics *vs*. Speech	Recog familiar	.011	.086	.897	-.158	.180
Familiar Lyrics *vs*. Unfamiliar Lyrics		.139	.086	.106	-.030	.308
Unfamiliar Lyrics *vs*. Melody		-.182	.086	.035*	-.351	-.013
Familiar Lyrics *vs*. Speech	Recog unfamiliar	.001	.107	.990	-.209	.211
Familiar Lyrics *vs*. Unfamiliar Lyrics		.139	.107	.194	-.071	.349
Unfamiliar Lyrics *vs*. Melody		-.310	.107	.004*	-.520	-.100

Note. * denotes significance.

4.3.2 Onset time.

For each participant onset time (OST) in seconds taken to begin production of the target lyrics was recorded and showed the shortest OST in quiet conditions (M = 2.18, SD = 1.22), a rise when accompanied by the unfamiliar distracter (M = 2.23, SD = 1.59), with a further rise during the familiar distracter condition (M = 2.38, SD = 1.68). A 3 (Sound Condition: quiet, familiar, unfamiliar) × 4 (Sound Type: Familiar Lyrics, Melody, Speech, Unfamiliar Lyrics) mixed ANOVA showed Sound Condition to be non-significant, $F(2, 568) = 2.438$, $MSE = 1.226$, $p = .088$, $\eta_p^2 = .009$). This indicates that time taken to begin speaking the target lyrics did not differ significantly according to the presence of a familiar or unfamiliar distracter as opposed to quiet, see Figure 7. The between-participants effect of Sound Type was non-significant, $F(3, 284) = .079$, $MSE = 4.439$, $p = .971$, $\eta_p^2 = .001$). There was also a non-significant Sound Condition × Sound Type interaction, $F(6, 568) = .645$, $MSE = 1.226$, $p = .694$, $\eta_p^2 = .007$). Therefore, OST to begin speaking the target rhyme was unaffected by the nature of the distracter, Familiar sung-lyrics, Melody, Speech (spoken-lyrics), or Unfamiliar sung-lyrics, as a function of its familiarity.

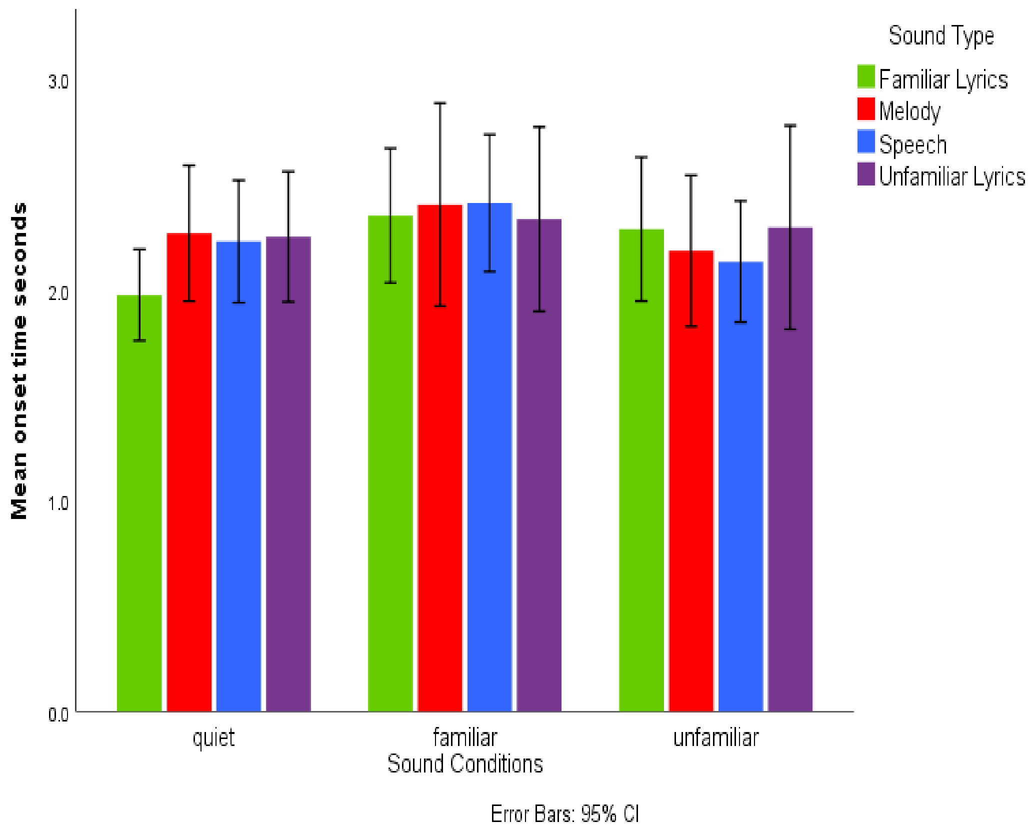


Figure 7. Mean onset time to begin speaking the target lyrics across Sound Conditions by Sound Type Group.

4.4 Discussion

The main purpose of Study II in this series was to address the question: Are melody and lyrics (within song) integrated or represented independently? Therefore, the aim was to examine differences in recall accuracy of familiar song lyrics from long-term memory by speaking performance. Sound Conditions, Sound Type, and assessment procedures were identical to those of Study I (Chapter III). Therefore, this discussion addresses issues regarding patterns of accuracy resulting from a retrieval process from long-term memory reliant on access to semantic/syntactic representations (e.g., Beaman, Hanczakowski, Hodgetts, Marsh, & Jones, 2013; Jones et al., 2012; Marsh et al., 2017).

4.4.1 Integration or independence?

Although melody and lyrics have been shown to be integrated in memory in consideration of the semantic aspects (Schön et al., 2010; Serafine et al., 1984), Study II results support a degree of independence (Besson et al., 1998, 2001; Peretz et al., 1994). For Study II, sung-lyrics, irrespective of familiarity, impaired target lyrics retrieval to a greater degree than spoken-lyrics, familiar or unfamiliar. Tentatively this could suggest that lyrics within song are stored differently to verbal information for long-term memory retrieval as has been shown to occur in WM imitation tasks (Berz, 1995; Besson & Schön, 2001; Halpern & Müllensiefen, 2008; Schulze et al., 2011). Moreover, the fact that unfamiliar spoken-lyrics, were more disruptive than unfamiliar melody also suggests different pathways and a degree of storage separation, as in Study I (melody retrieval) the opposite pattern of disruption was observed. Irrelevant melody (familiar or unfamiliar) was significantly more distracting than irrelevant spoken-lyrics for retrieval/production of a familiar melody.

Following the modular model of music processing (based on Peretz & Coltheart, 2003; Figure 4) concurrent presentation of irrelevant familiar lyrics, whilst attempting to produce target lyrics, generates greater activation in the phonological lexicon than an irrelevant familiar melody, that would proceed directly to the musical lexicon. However, some weak link between the musical and phonological lexicon would be generated by associated lyrics of the familiar melody thereby creating some impairment to target lyrics performance. The results of Study II show lyrics production more detrimentally affected by sung-lyrics than by spoken-lyrics or melody alone.

Based on the interaction of auditory pathways interference to motor control, perceptual processes, auditory-motor mapping, and memory, would inhibit vocal production of target lyrics (Dalla Bella et al., 2015). However, the interaction between semantic and musical memory, melody supporting lyrics retrieval, albeit an observation rather than an empirical finding, suggests an overlap of processing (Gröussard et al., 2010; Koelsch et al., 2002; Peynircioğlu et al., 1998).

4.4.2 Motor-planning and production.

For Study II there were two added components to routinely used serial recall experiments in the presence of irrelevant sound: retrieval from long-term memory, and production of target information by speaking. The skill of speech-planning, adopted to aid transition of sentence flow, requires the assembly of an articulatory plan. In this approach, the sequence of words, as opposed to individual words, are paramount, and this coheres with the perceptual-gestural view whereby the motor planning process operates according to the sequence of items rather than their individuality (e.g., Hughes & Jones, 2005). Consequently, the interplay between “streams” formed by the auditory input and the sequence formation of sub-vocal gestures is influential to task accuracy. Decreased performance in the presence of irrelevant sound containing lyrics for Study II may, therefore, reflect the efficacy of assembly and sub-vocal rehearsal of lyrics sequence patterns.

This finding, in part, supports the concept that the motor-planning mechanism underpinning speech is more vulnerable to sung-lyrics as compared to spoken-lyrics. Irrelevant familiar sung-lyrics would appear to have affected the motor-programming necessary to retrieve and perform familiar target lyrics (Leman, 2007; Leman & Maes, 2015), the vocal production vulnerable to pre-attentive serial processing of distracter lyrics/melody. It can be argued in the current study that the limited degree of change to successive sounds in the spoken-lyrics monotonic delivery—perceived from the same environmental origin, following the same stream—as compared to the more prosodic, undulating pitch contour of the sung-lyrics conditions—speculatively perceived from a different environmental origin—reduced conflict between the two processes of seriation (automatic for irrelevant sound and deliberate for target lyrics).

The center of motor-planning, the supplementary motor area (SMA; Tanaka & Kirino, 2017) integrates multi-modal information in musical imagery. During the current study, an integration of multi-modal information, such as sensory, cognitive, emotional, required for performance may have been recruited in addition to motor control to facilitate target lyrics speaking (Pfordresher, 2006; Tanaka & Kirino, 2017). It cannot be ruled out that the vocal-motor production necessary to rehearse and deliver the speaking output, from representation within long-term memory, during simultaneous presentation of musically semantically related (familiar) irrelevant auditory information may have been aided, or hindered, by auditory imagery (Bailes, 2007; Smith, Reisberg, & Wilson, 1992). Therefore, participants' auditory imagery ability may have been influential in how components of previously encoded songs were retrieved from their memory (Brown & Palmer, 2013).

Although, imagery ability when required to generate lyrics (use of the "inner-ear" or "inner-voice") has been identified in earworm studies (e.g., Beaman & Williams, 2010), the motor cortex reaction to music contributing to the earworm effect even when "listening" to music in one's head, the extent to which musical imagery was influential in Study II is unclear. It is feasible, however, that partnership between the "inner ear" and "inner voice" necessary for retaining verbal items may have been compromised (Smith et al., 1995). Participants' ability to imagine the associated familiar melody of the target rhyme lyrics may have aided or impaired vocal-motor production of the corresponding lyrics, or, conflicted with the irrelevant melody/associated lyrics.

4.4.3 Familiar melody *versus* unfamiliar melody comparison.

Given that across-experiment differences in the patterns of performance as a function of the distracter conditions would constitute support for the independence view, it would be useful to highlight consistencies and differences observed between Study I-II. The different effects produced in Study I (melody humming) and Study II (lyrics speaking) may best be explained first with reference to melody familiarity within each Sound Type Group. With regard to familiar *versus* unfamiliar irrelevant melody, Study I-II did not appear to be affected in the same way according to the familiarity of the melody. Although for the Familiar Lyrics Group a familiar melody is more disruptive to lyrics retrieval than an unfamiliar melody, this was not observed in the melody humming task. The incongruity of known lyrics to a <u>novel melody</u> created increased disruption when the focal task required speech planning assembly and production of lyrics. For the Melody Group, although familiar melody is more disruptive than unfamiliar melody in the lyrics speaking task, this is not shown for melody humming. In part this may have been due to the imagined lyrics of the irrelevant familiar melody generating increased competition with vocal-motor production required for speaking as compared to humming. The Speech Group showed no difference between humming, or lyrics retrieval, the absence of irrelevant lyrics overriding melody familiarity. With regard to the Unfamiliar Lyrics Group, the reverse pattern to the Familiar Lyrics Group presents. When combined with unfamiliar lyrics a familiar melody is more potent than an unfamiliar melody in melody humming but, although close, this did not observe for lyrics speaking. It appears that the incongruity created by novel lyrics to a familiar melody is more disruptive to melody humming than lyrics speaking.

With regard to Sound Condition, first, a familiar sound did produce some consistent patterns. For both melody and lyrics retrieval the Familiar Lyrics *versus* Melody Group, and Familiar Lyrics *versus* Speech Group were impaired significantly. However, for the Unfamiliar Lyrics Group neither humming or lyrics retrieval produced a significant effect. This again shows the potency of lyrics/melody incongruity when planning and executing vocal-motor planning production be that for speaking or humming. Second, there were also a number of inconsistencies between Study I-II. With a familiar melody, Melody distraction is more disruptive than Speech (spoken-lyrics) distraction when humming the target melody but not when lyrics speaking. This suggests a familiar melody has a stronger pathway than spoken-lyrics from the musical lexicon generating increased competition with vocal planning for target output. Speech *versus* Unfamiliar Lyrics Sound Type is significant for humming but not for lyrics retrieval. But for Melody *versus* Unfamiliar Lyrics Groups, a familiar sound is only significantly disruptive when speaking target lyrics.

With an unfamiliar Sound Condition some consistent patterns emerged. For both melody humming and lyrics speaking Familiar Lyrics *versus* Melody is significant, as is Melody *versus* Speech. However, with an unfamiliar distracter retrieval disruption differs. Whereas for Familiar Lyrics *versus* Speech Sound Types, significant effect emerged for melody humming but this did not observe for lyrics speaking. An inconsistent pattern also emerged for Familiar Lyrics *versus* Unfamiliar Lyrics Groups as melody humming significant but lyric speaking non-significant. The Speech *versus* Unfamiliar Lyrics Group was also significant for melody humming but non-significant for lyrics speaking. However, Unfamiliar Lyrics *versus* Melody Groups showed the opposite pattern, a significant effect for lyric retrieval but not for melody humming.

4.5 Task comparison: Humming and semantic retrieval experiments

To assess more directly whether effects resulting from the two components of song identified in Study I (melody via humming) and Study II (lyrics via speaking) were as a consequence of experimental design, data from Study I, was compared with data from Study II. The main aims for this comparison were to: identify if patterns of disruption were similar or dissimilar across the two experiments (Task); identify if patterns of retrieval were affected by the Sound Condition, familiar as compared to unfamiliar, or by the addition of lyrics; ascertain if Sound Type contributed to similar or dissimilar distraction susceptibility. Retrieval performance from long-term memory in the presence of the same irrelevant sound was required for both Tasks, with no variation to procedure.

Baseline accuracy was relatively consistent across the humming and speaking tasks in the quiet condition (although slightly higher for the speaking task), but both tasks showed decreased accuracy in the presence of an unfamiliar distracter and a further decrease in the presence of a familiar distracter. However, interestingly decremental outcomes followed a relatively consistent pattern albeit with a broadening of the gap between tasks (Figure 8). These results indicate that across the melody humming and semantic speaking tasks, the distracter condition was influential in accuracy of performance irrespective of familiarity of Sound Condition.

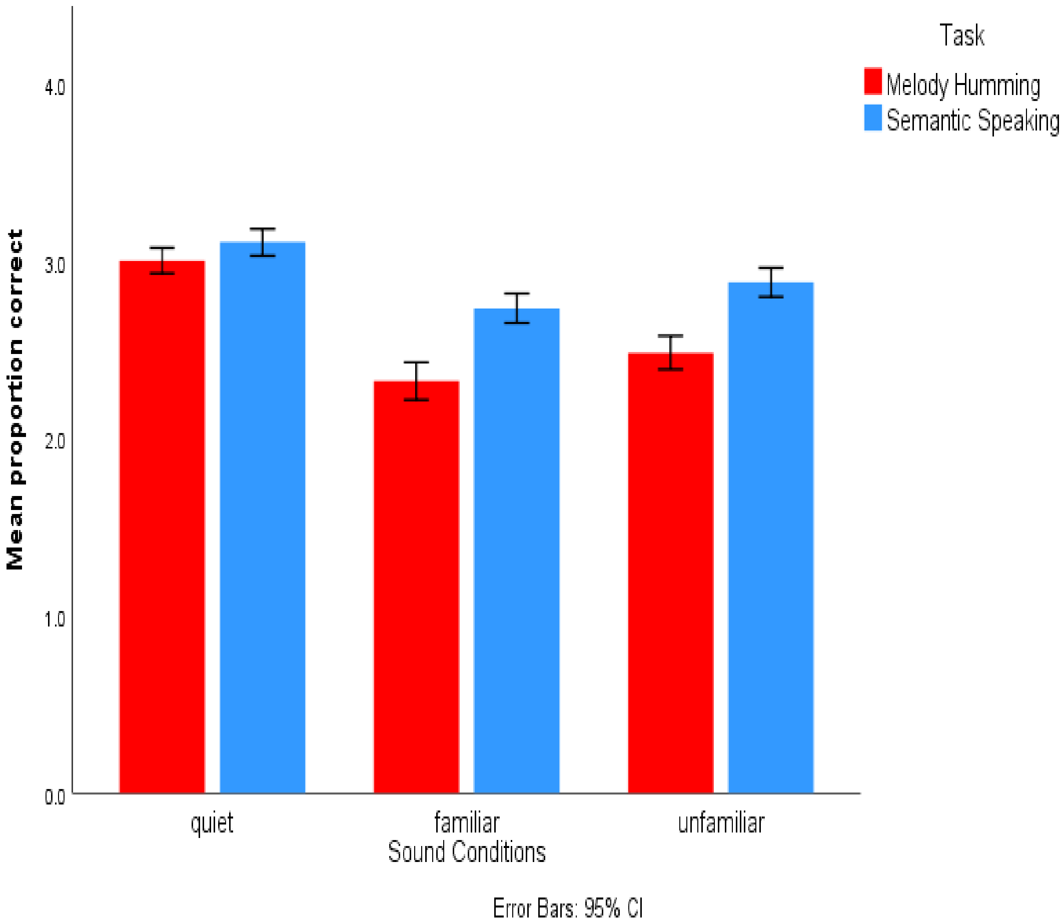


Figure 8. Retrieval score means across the measure of Sound Condition according to Task.

A mixed ANOVA incorporated a 3 (Sound Condition: quiet, familiar, and unfamiliar) × 2 (Task: Humming and Semantic) × 4 (Sound Type: Familiar Lyrics, Melody, Speech, Unfamiliar Lyrics) factors. As expected, results identified a significant effect of Sound[36], $F(2.000, 1148.000) = 203.863$, $MSE = .207$, $p < .001$, $\eta_p^2 = .262$. There was also a significant between-participants effect of Task, $F(1, 574) = 32.803$, $MSE = 1.187$, $p < .001$, $\eta_p^2 = .054$, and Sound Type, $F(3, 574) = 8.269$, $MSE = 1.187$, $p < .001$, $\eta_p^2 = .044$. These results validate the preceding analysis (Chapter III, 3.3 and Chapter IV, 4.3).

The focus of this comparison analysis is concerned with any interaction of Task with other factors. There was a significant interaction between factors: Sound Condition × Task, and a 3-way interaction between Sound Condition × Task × Sound Type. In addition, the between-participants Task × Sound Type interaction was significant as shown in Table 4.8.

Table 4.8

Results from a 3 (Sound Condition) × 2 (Task) × 4 (Sound Type) Mixed ANOVA

Within-participants factors	df	*F*	*MSE*	*p*	η_p^2
Sound Condition × Task	2, 1148	20.279	.207	< .001	.034
Sound Condition × Sound Type	6, 1148	8.964	.207	< .001	.045
Sound Condition × Task × Sound Type	6, 1148	4.731	.207	< .001	.024
Between-participants factors					
Task × Sound Type	3, 574	3.300	1.187	.020	.017

Pairwise comparisons for Task confirmed significantly lower recall scores for Task 1, melody humming, as compared to Task 2, lyric speaking ($MD = -.299$, $SE = .052$, $p < .001$; 95% CI [-.401, -.196]). A simple effects analysis showed this to be driven by the Familiar Lyrics, and the Melody Sound Type Groups across both Tasks, see Table 4.9.

Table 4.9

Humming task (Hum) and Semantic task (Sem) comparison according to Sound Type

Task Hum *vs.* Sem	Sound Type Group	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
	Familiar Lyrics	-.407	.103	< .001	-.608	-.205
	Melody	-.510	.105	< .001	-.716	-.304
	Speech	-.116	.105	.268	-.322	.090
	Unfamiliar Lyrics	-.162	.105	.123	-.368	.044

[36] Huynh-Feldt reported following sphericity violation $\varepsilon = 1.000$.

A simple effects analysis to decompose the 3-way interaction Sound Condition × Task × Sound Type comparing Task, identified only the non-music Speech Group to be non-significant across all conditions. All other Sound Type Groups were significant with a familiar and unfamiliar distracter, see Table 4.10. In addition, the Melody Group was significant when comparing Task during the quiet condition, suggesting a more knowlegable Melody Group of participants for the semantic Task.

Table 4.10

Comparative Humming task (H) and Semantic task (S) three-way interaction: Task × Sound Type × Sound Condition

Task	Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
H *vs.* S	Familiar Lyrics	familiar	-.500	.131	< .001	-.757	-.242
		unfamiliar	-.610	.121	< .001	-.847	-.372
H *vs.* S	Melody	quiet	-.290	.107	.007	-.499	-.081
		familiar	.578	.134	< .001	-.840	-.315
		unfamiliar	-.663	.123	< .001	-.905	-.420
H *vs.* S	Unfamiliar Lyrics	familiar	-.424	.134	.002	-.686	-.161

It is clear from the comparison analysis that there are some similar significant effects between Task (Study I-II) when comparing Sound Type Group, see Table 4.11. However, some dissimilar effects arose as shown in Table 4.12. Of particular interest from the three-way interaction comparing Sound Type Group is the significant effect obtained in the Humming task for the Melody Group *versus* the Speech Group with a familiar Sound Condition (*MD* = -.296, *SE* = .134, *p* = .027; 95% CI [-.558, -.033]) that did not obtain for the Semantic task (*MD* = .167, *SE* = .134, *p* = .213; 95% CI [-.069, .429]). However, with an unfamiliar Sound Condition, a dissociation arose. In the Humming task, a significant effect occurred between Melody and Speech Sound Type Groups (*MD* = -.299, *SE* = .123, *p* = .016; 95% CI [-.541, -.056]) indicating humming a target melody was disrupted more by an unfamilar distracter melody than by unfamiliar spoken-lyrics. For the Semantic task a significant effect was identified between the Melody and Speech Groups (*MD* = .299, *SE* = .123, *p* = .016; 95% CI [.056, .541]), indicating that lyrics production performance was disrupted more by an unfamiliar spoken-lyrics distracter than an unfamiliar distracter melody. Table 4.13 shows a comparison of familiar *versus* unfamiliar Sound Condition according to Task and Sound Type Group.

Table 4.11

Task comparison according to Sound Condition and Group showing similar effects

Task	Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Hum	Familiar Lyrics *vs.* Melody	familiar	-.361	.131	.006	-.619	-.103
Sem			-.439	.134	< .001	-.702	-.176
Hum	Familiar Lyrics *vs*. Speech		-.657	.131	< .001	-.914	-.399
Sem			-.272	.134	.042	-.535	-.010
Hum	Familiar Lyrics *vs.* Melody	unfamiliar	-.417	.121	<.001	-.654	-.179
Sem			-.469	.123	< .001	-.712	-.227
Hum	Melody *vs.* Speech		-.299	.123	.016	-.541	-.056
Sem			.299	.123	.016	.056	.541

Table 4.12

Task comparison according to Sound Condition and Group showing dissimilar effects.

Task	Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Hum	Melody *vs.* Speech	familiar	-.296	.134	.027*	-.558	-.033
Sem			.167	.134	.213	-.096	.429
Hum	Speech *vs.* Unfamiliar Lyrics		.471	.134	< .001*	.208	.733
Sem			.163	.134	.225	-.100	.425
Hum	Speech *vs.* Unfamiliar Lyrics	unfamiliar	.274	.123	<.001*	-.362	.056
Sem			.119	.123	.333	-.123	.362

Note. Hum - Humming Task; Sem - Semantic Task.

Table 4.13

Task comparison of familiar vs. unfamiliar Sound Condition according to Group

Task	Sound Type Group	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Hum	Familiar Lyrics	familiar *vs.* unfamiliar	-.069	.073	.340	-.212	.073
Sem			-.179	.075	.018*	-.327	-.031
Hum	Melody	familiar *vs.* unfamiliar	-.125	.075	.098	-.273	-.023
Sem			-.210	.075	.006*	-.358	-.061
Hum	Speech	familiar *vs.* unfamiliar	-.128	.075	.091	-.276	.020
Sem			-.078	.075	.303	-.226	.070
Hum	Unfamiliar Lyrics	familiar *vs.* unfamiliar	-.325	.075	<.001*	-.473	-.177
Sem			.121	.075	.110	-.269	.027

Note. * denotes significance.

A comparison of onset-time (OST) in seconds taken to begin production of the required target according to task showed the shortest OST across the three Sound Conditions for the semantic task (lyric speaking). However, OST to begin humming the target melody or speaking the target rhyme was not significantly affected by the nature of the task (p = .148). Therefore, the OST data time and results are not reported here.

4.6 Discussion

Comparative data analysis aimed to identify: similarities and differences in patterns of disruption across the Task employed in Study I-II; if patterns of retrieval were differentially affected by familiar as compared to unfamiliar sound, or sung-lyrics as compared to spoken-lyrics, and to determine distraction susceptibility according to Sound Type.

4.6.1 Overall results.

The key finding from the comparison analysis indicates unequivocal significance between Study I and Study II. Study I, melody retrieval via humming performance, was impaired to a greater degree than Study II, lyrics retrieval via speaking, with the same irrelevant auditory distracters. Furthermore, accuracy of performance was reduced across all auditory conditions, for both studies, irrespective of familiarity of the irrelevant sound. Both studies showed a familiar distracter to impair performance to a greater extent than an unfamiliar distracter, although this was not always significant. OST was significantly reduced by irrelevant sung-lyrics, familiar or unfamiliar, for melody humming, but Sound Type did not differentially impact on time taken to begin lyrics speaking.

4.6.2 Sound Condition × Task × Sound Type.

A three-way interaction between Sound Condition × Task × Sound Type clearly showed participants' ability to produce a target melody (humming, Study I) or target lyrics (speaking, Study II) differed in the presence of irrelevant speech irrespective of familiarity. Furthermore, regardless of the task undertaken, performance accompanied by an unfamiliar Sound Condition was unaffected. However, during an irrelevant familiar melody Sound Condition, participants' retrieval in the melody and all sung-lyrics groups were more greatly impaired in the humming Task (Study I) as compared to the speaking Task (Study II). For example, lower scores were obtained when humming a familiar target melody (e.g., 'Mary, Mary, quite contrary'), in the presence of another familiar melody (e.g., 'London Bridge', Figure 1a[37]) than speaking the lyrics. It was also more difficult to hum the target melody (e.g., 'Mary, Mary, quite contrary') in the presence of irrelevant familiar sung-lyrics (e.g., 'London Bridge', Figure 1c) and irrelevant unfamiliar sung-lyrics (e.g., 'Hold my hand', Figure 1e).

Interestingly, during an unfamiliar melody Sound Condition the Unfamiliar Lyrics Group Tasks were not differentially impaired. In other words, the unfamiliar lyrics coupled with an unfamiliar melody (e.g., 'Hold my hand', Figure 1f) appeared equipotent in distraction for humming the melody or speaking the lyrics of the target rhyme (e.g., 'Mary, Mary quite contrary'). But, in the Familiar Lyrics Group and the Melody Group participants were more impaired when humming (Study I) as compared to speaking (Study II). In other words, trying to hum the target melody (e.g., 'Mary, Mary, quite contrary') was more difficult than speaking the target words in the presence of unfamiliar melody (e.g., 'Hold my hand') combined with familiar lyrics (e.g., 'London Bridge', Figure 1d) or in the presence of an unfamiliar melody (e.g., 'Hold my hand', Figure 1b). It should be noted that in quiet conditions the Melody Group showed lower scores in the humming, as compared to the semantic, task identifying them as a more able cohort in Study II, see Table 4.10.

[37] Distracter examples Figure 1a-h shown in Chapter III 3.2.3.

4.7 Comparison of overall results with previous findings

Routinely, the presence of task irrelevant sound has been found detrimental to recall task accuracy. However, similarity in process, rather than content, between two contemporaneous streams appear to dictate the magnitude of the disruption irrelevant information confers to focal task activity (e.g., Jones et al., 1992; Marsh et al., 2009; Salamé & Baddeley 1982). For Study II, following the establishment of a similar baseline in quiet conditions across the four between-participant Sound Type Groups, Familiar Lyrics, Melody, Speech, and Unfamiliar Lyrics (although the Melody Group was the highest scoring and significantly more able than the Unfamiliar Lyrics Group), all irrelevant sound, with or without lyrics, disrupted lyrics retrieval from long-term memory compared to quiet. Although results from this present novel task show a decrease in ability to retrieve lyrics of target nursery rhymes when subjected to any of the distracter conditions, speaking of lyrics was impaired less by irrelevant instrumental melody, familiar or unfamiliar, than by irrelevant spoken-lyrics, or, to a greater degree, sung-lyrics. The appreciable increased impairment to retrieval by the presence of irrelevant sung-lyrics lends support to previous literature identifying vocal, as opposed to instrumental, music to be more detrimental to task accuracy (e.g., Alley & Greene, 2008; McCorkell, 2012; Nittono, 1997; Perham & Vizard, 2011; Salamé & Baddeley 1989).

4.7.1 Integration or independence?

The comparison between the retrieval of known melodies, from long-term memory via humming (Study I), and the semantic retrieval of known lyrics, from long-term memory (Study II), clearly shows that the melody humming was impaired to a greater degree than lyrics speaking with the same irrelevant auditory distracters. This implies that there is some degree of processing separation. If melody and semantic processing were undertaken within the same processing system or set of shared processes, then accuracy for the melody and lyrics retrieval would have been comparable. Irrelevant background sounds impaired performance across all sound conditions across both studies. But, different patterns of disruption occurred according to task. Study I-II clearly show task accuracy decreases with the same stimulus. In other words, humming a melody was less accurate when accompanied by a melody distracter than by spoken- or sung-lyrics. In addition, the retrieval of song lyrics was less accurate when accompanied by irrelevant spoken- or sung-lyrics than melody alone (even when associated, and probably encoded, with lyrics). This alone suggests some degree of overlap in the way speech and tones are processed (Williamson et al., 2010).

Furthermore, findings would relate to the "inner ear" account (Macken & Jones, 1995; Smith et al, 1995). Subvocal processes were engaged through both tasks to aid/rehearse the corresponding required retrieval against competition generated from irrelevant sound stimuli or imagined from a visual title displayed through performance.

The results of the current study, on the whole, gel neatly with the modular model of music processing (Peretz & Coltheart, 2003; Figure 4). According to the notion of music *versus* lyrics processing independence offered by this account, the patterns of the results across Study I-II are generally explicable through the assumptions of the interplay between the phonological lexicon and the music lexicon, the music lexicon and vocal plan formation and the phonological lexicon and spoken output. The greater disruption produced by spoken-lyrics against melody is due to the task-irrelevant material activating irrelevant entries in the phonological lexicon that impair spoken output of lyrics. Similarly, the greater disruption of melody against spoken lyrics in the context of humming is due to the task-irrelevant melody causing activation to flood into the vocal plan formation for melody retrieval. Familiar melody might be more distracting than unfamiliar melody in the case of lyrics production due to the bidirectional links between the music lexicon and the phonological lexicon, causing activation of lyrics within the phonological lexicon, thereby competing with the target information, as cued by song-title, for retrieval. The familiarity effect within the context of melody is arguably more important in lyric, than humming, retrieval because in the case of the latter, unfamiliar melody can still affect the vocal plan formation for melody production but would be expected to have much less influence on spoken output of lyrics (e.g., Jones et al., 2012). For production of spoken-lyrics, the combination of unfamiliar lyrics with unfamiliar melody should produce less activation within the phonological lexicon than a combination of familiar lyrics with familiar melody because in the latter case there is priming of the lyrics in the phonological lexicon from melody, in addition to activation from the linguistic input. However, only activation of lyrics within the phonological lexicon would be expected from the unfamiliar lyric component of the unfamiliar melody-unfamiliar lyrics combination. In the case of humming, the combination of familiar lyrics with familiar melody would be more disruptive because the familiar melody strongly activates a competing vocal plan for melody recall.

Humming performance is impaired by unfamiliar lyrics coupled with a familiar melody because the familiar melody would activate a competing vocal plan for production of a melody. However, when lyrics production is required, the disruption produced to spoken-lyrics retrieval by familiar melody is inhibited by the activation of the unfamiliar lyrics that impair speech output-planning of lyrics. In the case of spoken distracters, a familiarity effect would be unlikely to occur because it might be assumed that there is an asymmetry in the bidirectional connection between music and phonological lexicons to the extent that spoken familiar lyrics do not activate familiar melody to the same extent than the reverse. As such, competition incurred from the mere activation of lyrics within the phonological lexicon is not exacerbated through incrementing activation caused via input from the musical lexicon. Such an explanation would also account for why familiar spoken-lyrics fail to produce greater disruption than unfamiliar spoken-lyrics in the case of humming—the familiar spoken lyrics producing only weak activation of the associated melody that would impair the vocal plan formation required. This overarching account appears plausible but requires refining, particularly from studies investigating the strength of bidirectional priming between lyrics and melody (cf. Peretz, Radeu, & Arguin, 2004).

4.7.2 Interference-by-process *versus* interference-by-content.

Study I explored whether the disruption to humming performance was attributable to interference in the content of the task-irrelevant material or due to the task-irrelevant material competing for the same process as the focal task material. It was argued that the pattern of data were more consistent with the interference-by-process account subsumed within the perceptual-gestural framework (Jones et al., 2004, 2006). A key construct of the interference-by-content account is that similarity between the representations within the same memory space or store determines disruption (Neath, 2000; Salamé & Baddeley, 1982). Inadequacies of the interference-by-content account are underscored by several findings in the context of the current study. For example, familiar melody produced more disruption than unfamiliar melody despite little overlap in the similarity in content between the task-irrelevant and to-be-recalled material. One possible explanation for the difference in disruption, in line with the interference-by-content view, is that familiar melody automatically activates familiar lyrics and these disrupt retrieval of target lyrics (Pring & Walker, 1994). However, the view does not seem to adequately explain why familiar lyrics combined with familiar melody would be more disruptive to lyric retrieval than unfamiliar lyrics combined with familiar melody since in both cases the presence of task-irrelevant lyrics should interfere with the representation and production of task-relevant lyrics. This is since, according to the interference-by-content account, familiarity should not be important. Furthermore, the interference-by-content account does not adequately explain how performance disruption attributable to different properties of task-irrelevant sound are acutely sensitive to the nature of the focal task processing.

The process-oriented view held by the perceptual-gestural account and embodied in the interference-by-process concept, readily explains the differential susceptibility of performance to different properties of sound and holds that the nature of the disruption observed is predicated on the processes underpinning the focal task. Within this account, changing acoustic properties of the irrelevant sound and the demands of seriation within the focal cognitive task dictate disruption in tasks requiring the production of a sequence (Jones et al., 2004, Jones & Tremblay, 2000). In Study I it was held that competition between the seriation of a task-irrelevant sequence and a target sequence dictated disruption with the competition stemming in part from activation of task-irrelevant melodies within the musical lexicon. Generally melodic information produced greater disruption because it superficially (and more specifically in the case of familiar excerpts) fits the generic skill of producing a target melody through motor processing. In Study II, the requirement to produce a sequence was more governed via semantic, than melodic retrieval. In this case, the lexical information activated by task-irrelevant sound, over melodic information, became the more potent distracter since it coheres more closely with the response criteria—to produce lyrics. Melodic information becomes a less potent distracter since the requirement for seriation within the output process is much diminished. In part, the additional disruption produced by familiarity could be due to stronger competition for retrieval and the legacy of dynamic mechanisms recruited to prevent that irrelevant information from assuming the control of action, or removing it from the control of action should it succeed in doing so (see Chapter I for further discussion).

4.7.3 Interference-by-process *versus* attentional capture.

In the current study, the interference-by-process view is favoured to explain the different patterns of disruption to lyric retrieval in the presence of task-irrelevant sound with different characteristics. An alternative account is that the disruption produced by task-irrelevant sound is attributable to attentional capture. The classical attentional capture view assumes that the mere presence of task-irrelevant sound captures attention away from any cognitively demanding focal task (e.g., Bell et al., 2010). This account, however, does not explain why familiar task-irrelevant auditory sequences would produce any more attentional capture than unfamiliar sequences. The problem with this account, as for the results reported in Study I, is that there is no material acoustic difference (e.g., presence of greater changing-state information, or unexpected events) between the familiar and unfamiliar sequences that would bestow on them differentially disruptive power. Further, the account cannot explain why disruption produced by various features of the task-irrelevant sequences are so peculiarly sensitive to the nature of the focal task itself. For example, the attentional capture account, as characterised by the notion of an orienting response (Cowan, 1995), cannot explain why lexical information, as compared with melody, is a more potent distracter when the focal task requires spoken retrieval of lyrics as compared with retrieval of a melody by humming. Similarly, the account does not explain why the reverse is true: that task irrelevant melody is more disruptive than lexical information, when the task requires hummed melody retrieval against retrieval of spoken-lyrics. To the contrary, the importance of similarity in processing between task-irrelevant and task-relevant sources of information is a central tenet of the interference-by-process view (Jones & Tremblay, 2000; Marsh et al., 2009). In Study III-IV variants of the attentional capture view will be further addressed (Hughes, 2014) and related to patterns of disruption observed across tasks that involve different processing demands.

4.8 Musicians *versus* non-musicians

Study I showed that no significant advantage or disadvantage emerged for performance in the presence of distracters when level of musical training among participants was considered. This finding stands in contrast to that of Thompson and Yankeelov (2012) who found a disadvantage for musicians on music tasks with music interference, as compared to language interference, unlike non-musicians who experienced the same degree of impairment. However, for Study II a significant advantage was observed following 8+ years of musical training or current participation in a choir (see 4.2.1). This may suggest enhanced retention of nursery rhyme lyrics routinely used in early music making. It also adds some support for short-term recall studies using music that have identified superior recall for musically trained participants (Elliott & Cowan, 2005). This finding runs contrary to expectation but supports the idea that advantage of musical training may be task-dependent.

4.9 Conclusion and subsequent research

In summary, performance in Study I-II was detrimentally affected by all Sound Conditions as compared to quiet, across all Sound Type Groups. In addition, for both tasks unfamiliar irrelevant sound impaired performance less than familiar irrelevant sound, the level of disruption qualified by group. The key finding was that retrieval accuracy for Study I and Study II did not show similar patterns of retrieval in the presence of the same Sound Type distracters. This points to some independence in melody/lyrics processing (Besson et al., 1998; Bonnel et al., 2001; Peretz, 1994). It also lends support for the strong pathways through the modular model of music processing (Peretz & Coltheart, 2003) from input to output for familiar as compared to unfamiliar musical material, with greater potential to impact on vocal-motor production planning and output. The production of melody/lyrics rather than their recognition, necessitated an articulatory plan to be formed and in so doing was reliant of the unfolding sequence of pitch/words rather than their individuality. This relates to the perceptual-gestural view (Hughes & Jones, 2005). Results also support the concept that the motor-planning mechanism underpinning speech is more vulnerable to sung-lyrics as compared to spoken-lyrics. Irrelevant familiar sung-lyrics appear to affect the motor-programming necessary to retrieve and perform familiar target lyrics (Leman, 2007; Leman & Maes, 2015).

Of interest also was the significant difference between humming (Study I) and semantic (Study II) retrieval, with increased scores obtained across all Sound Conditions, and Sound Types, for Study II. However, it is unclear if the lower scores obtained across all Sound Conditions and Sound Type in Study I, compared to Study II, was the result of an impairment in ability to retrieve the target or to perform in the presence of another familiar melody/lyrics, audible, or generated through imagining and subvocalizing the lyrics (Pring & Walker, 1994). For Study I-II seriation required for task performance was not, arguably, as obvious as for serial recall tasks. Therefore, it would now be beneficial to compare findings from Study I-II, with a task performed under the same irrelevant sound conditions but involving strict serial recall. The central purpose of the forthcoming studies, therefore, will be to address the attentional capture accounts and pitch them against the perceptual gestural view. Findings from a short-term memory task requiring item recall in order of presentation may cast light as to whether the scores garnered from Study I-II can be replicated with a non-music task (as melody and associated lyrics retrieval were both components of familiar song from long-term memory) and so determine if these results are task-process specific, or task-process independent. Subsequent results would further add to the debate regarding the familiarity of the distracter sounds causing attentional capture or interference-by-process.

CHAPTER V

EMPIRICAL STUDY III: DISRUPTION OF A SERIAL RECALL TASK BY IRRELEVANT SOUND

Abstract

Although task-irrelevant sound can produce considerable disruption of visually based tasks, particularly those tapping short-term memory for serial recall (e.g., Colle & Welsh, 1976; Jones, 1993; LeCompte, 1994; Salamé & Baddeley, 1982), the underlying mechanisms for such effects remain contested. Retrieval of melody (Study I) or lyrics (Study II) was primarily disrupted by familiar melody accompanied by either familiar or unfamiliar lyrics. While spoken-lyrics disrupted performance more than melody in Study II, the familiarity of lyrics or melody alone did not influence distraction. These results were interpreted as illuminating a specific interference-by-process whereby disruption is attributable to a conflict between deliberate and automatic processes of seriation (Jones & Tremblay, 2000). Familiar sung-lyrics capable of interfering with vocal-motor processing necessary for target melody (Study I) or lyrics (Study II) production as identified in Chapters III-IV. An alternative approach supposes task disruption is due to sound (e.g., familiar sung-lyrics) capturing attention away from the focal task (Bell et el., 2012; Cowan, 1995). If so, then such attentional capture should be task-independent. To address this, Study III explored whether the same patterns of disruption from the auditory distracter conditions adopted for Study I-II would manifest for a visual-verbal short-term memory serial recall task that, like melody recall (Study I) and lyrics recall (Study II), requires vocal-motor processing— subvocalization—for serial rehearsal of series of eight digits, but, unlike Study I-II does not require access to, or production of, melody/lyrics. On the interference-by-process view, familiar, as compared to unfamiliar, song should not produce greater disruption to serial recall, whereas with the attentional capture account familiar song should be the more disruptive. Consistent with the interference-by-process view, and at odds with the attentional capture view, the combination of familiar melody and lyrics was <u>not</u> a potent cause of disruption in serial recall. Thus, the nature of irrelevant sound effects can be dissociated by task requirements. Surprisingly, however, familiar melody produced significantly more disruption than unfamiliar melody, a pattern observed in Study II but not in Study I. Results are discussed in terms of interference-by-process (e.g., Jones & Tremblay, 2000) and attentional capture accounts (Bell et al., 2012; Cowan, 1995).

5.1 Introduction

Disruption to visual-verbal short-term memory performance resulting from the presence of irrelevant sound is referred to as the classical irrelevant sound effect (ISE; Beaman & Jones, 1997) and has been widely associated with serial recall tasks since only minimal disruption is observed when seriation is not required (Beaman & Jones, 1997; Klatte, Meis, Sukowski, & Schick, 2007). The ISE is a robust finding, consistently observed and evidenced across a wide range of studies (e.g., Colle & Welsh, 1976; Hughes et al., 2005, 2007; Jones & Machen, 1993; Morris et al., 1989; Nittono, 1997). A number of findings support the suggestion that the ISE derives from a conflict of seriation processing between to-be-remembered and to-be-ignored items on serial recall tasks (e.g., Beaman & Jones, 1997, 1998; Hughes, 2014). A leading account of the ISE, interference-by-process (e.g., Jones & Tremblay, 2000), holds that task-irrelevant sounds gain automatic access to a representational system wherein they are organised into coherent streams. The product of this organisation flows directly into the articulatory-planning process (Jones et al., 1993). A similar mechanistic explanation for disruption of familiar melody retrieval by the mere presence of irrelevant familiar melody was proposed for Study I. Here, it was suggested that retrieval of familiar melody was impaired by the automatic processing of irrelevant familiar melody because task-irrelevant material assumed control, or threatened to assume control, of action of the sub-vocal motor system responsible for target melody production (Godøy & Leman, 2009; Leman, 2007; Leman & Maes, 2014, 2015). Despite the persuasiveness of the interference-by-process account (e.g., Jones & Tremblay, 2000), an alternative mechanism for disruption to focal task performance by task-irrelevant sound is attentional capture (e.g., Bell et al., 2012; Cowan, 1995).

There is an active and lively debate, within the serial short-term memory literature, between proponents of a unitary account of distraction (Bell et al., 2012; Elliott, 2002) and a duplex account (Hughes et al., 2007) of distraction. The former account posits that all forms of distraction are attributable to attentional capture and the latter account assumes auditory distraction can arise from either interference-by-process or an attentional capture mechanism (e.g., Hughes et al., 2007; Vachon et al., 2012). This current chapter aims to determine whether disruption of melody and lyrics retrieval by the presence of familiar melody accompanied by familiar or unfamiliar lyrics (over and above unfamiliar melody combined with unfamiliar lyrics) was produced in Study I-II due to active engagement in a melody and lyrics retrieval process—as the interference-by-process view argues. Or, whether it occurs because the combination of familiar melody with familiar or unfamiliar lyrics simply captures attention relative to other sound conditions. To this end, the potential disruptive effect of auditory distracters used in Study I-II were examined in the context of a serial recall task that, like melody and lyrics retrieval, requires vocal-motor process (Hughes & Marsh, 2017), but does not require retrieval of melody or lyrics. Specifically, this chapter addresses whether a dissociation in susceptibility to task performance by disruption via different properties of irrelevant sound (e.g., familiarity of melody/lyrics) can be modulated by the nature of goal-driven processes adopted (e.g., retrieval of melody/lyrics *versus* serial rehearsal), as the interference-by-process account suggests. Or, alternatively, whether disruption produced by different characteristics of irrelevant sound are insensitive to goal-driven process type, as suggested by the attentional capture account. The following subsections in this chapter address the main theories pertaining to mechanisms of distraction within the serial recall paradigm, a serial-recall task (Study III), with subsequent analysis and discussion.

5.2 The serial recall paradigm

Experimental studies on auditory distraction have relied heavily on the serial recall paradigm (e.g., Colle & Welsh, 1976; Hughes et al., 2005, 2007; Jones et al., 1992; Jones & Macken, 1993; Salamé & Baddeley, 1982) in which 7-9 digits or consonants are visually-presented in a sequence and participants recall these in serial order. Order report is required either immediately after visual presentation or following a brief retention interval of 5-10 seconds and involves rehearsal of presented items that utilizes the vocal system in a form of subvocalization to cohere a disparate sequence into a motor-plan for future action recall (Hughes et al., 2009). The ISE refers to the finding that the presence of to-be-ignored sound demonstrating appreciable acoustic variation produces considerable disruption of serial recall performance when coinciding with the presentation of visual items or filling a retention interval after the last to-be-remembered item and onset of a recall cue (Jones, 1993).

5.2.1 The duplex-mechanism account.

Auditory distraction type I: Interference-by-process.

The duplex-mechanism account of auditory distraction (Hughes, 2014) holds that distraction can arise through two mechanisms—interference-by-process and attentional capture—and thus there are two discrete types of auditory distraction. Interference-by-process incorporates the notion that task-irrelevant sound can disrupt cognitive task performance by interfering with particular processes involved in a given focal task (e.g., Jones & Macken, 1993; Jones & Tremblay, 2000). This mechanism was first established in the domain of short-term memory wherein serial recall is impaired by the mere presence of sound. A key feature of the interference-by-process account is susceptibility of serial recall tasks to the changing-state effect whereby an acoustic stimulus that exhibits appreciable acoustic variation over its time course (e.g., C, T, G, K) impairs serial recall performance more than a sequence of steady-state repeated sounds (e.g., C, C, C, C). The changing-state effect has been replicated by a multitude of serial recall studies within which changing-state characteristics are imbedded in noise (Furnham & Strbac, 2002; Perham et al., 2013; Salamé & Baddeley, 1989), sequences of tones (Jones & Macken, 1993; Tremblay & Jones, 1998), music (Ellermeier & Hellbrück, 1998; Nittono, 1997), song (Alley & Greene, 2008; Iwanaga & Ito, 2002; McCorkell, 2012), speech (Schlittmeier & Hellbrück, 2009), when recalling visually-presented digits (Pring & Walker, 1994), consonants (Williamson et al., 2010), or words (McCorkell, 2012). For example, staccato instrumental music containing distinct temporal-spectral variations (hence more changing-state) is more disruptive of serial recall than legato music, which comprises limited spectral variation (Schlittmeier & Hellbrück, 2009, Experiment 1). This is consistent with interference-by-process and the changing-state effect.

According to the interference-by-process account, changes in successive sound elements within a to-be-ignored sound yield cues as to the order of those sounds via the pre-attentive process of streaming (Bregman, 1990). This order information then interferes with the deliberate goal-driven process of serially rehearsing visual to-be-remembered items in preparation for sequential output via co-opting vocal-motor apparatus supporting covert speech (e.g., Jones & Macken, 1993). Steady-state sequences that comprise repeated items or continuous sound contain little, or no, element-to-element acoustic changes and thus yield very few order cues thereby producing minimal, if any, disruption of serial recall (Jones et al., 1992). Crucially, the interference-by-process account suggests that the changing-state effect is a joint product of processes applied to the sound, that are pre-attentive and involuntary, and the nature of conscious, deliberate processes supporting performance on the memory task (e.g., Jones et al., 1999; Jones & Macken, 1993; Macken et al., 2009).

While a thorough review is beyond the scope of this current chapter (for more detail see Chapter II), one pertinent finding is the non-monotonic relationship between the degree of change difference between successive elements within irrelevant sound and the degree to which it disrupts serial recall performance. For example, Jones et al.

(1999) demonstrated that increasing frequencies of successive tones from 0 to 2 to 5 semitones increased the magnitude of disruption up until separation distance between successive tones was 10 semitones where a point of fission occurs. Providing successive change increase within confines of fusion, order information is more strongly encoded the greater magnitude of change between successive elements. However, if change between successive items is too great, stream segregation will occur (Bregman, 1990; Warren & Obusek, 1972). This allows higher and lower pitched tones to form separate streams within which order information is lacking (the two streams are effectively steady-state). Consequent disruption to serial recall is minimal (Jones et al., 1999). Amenability of the ISE to modulation by streaming has also been shown by varying the spatial location of to-be-ignored sound. When all sound elements arise from the same spatial location, they form a single percept, or stream, wherein order information is strongly represented. However, assigning different items to different presentation locations results in the formation of different streams wherein order information is reduced. An ISE is thus greater in magnitude when sound items are delivered from the same spatial location (e.g., from a common carrier; Jones & Macken, 1995; Jones et al., 1999).

Since interference-by-process results from a conflict of seriation processes, the degree of disruption to a focal task rests upon the degree of seriation necessary to perform the focal task (Jones & Tremblay, 2000). Tasks requiring some degree of serial rehearsal are particularly susceptible to disruption via the changing-state effect (e.g., prose recall and mental arithmetic: Banbury & Berry, 1997; free recall: Beaman & Jones, 1997), while tasks that do not require such seriation are relatively immune to the changing-state effect (e.g., proof-reading: Jones et al., 1990; reading comprehension: Martin et al., 1988: missing-item task; Beaman & Jones, 1997; free recall of semantic category-exemplars: Marsh et al., 2008, 2009; and perceptual tests: Burani, Vallar, & Bottini, 1991). It should be noted, however, that the interference-by-process account (Jones & Tremblay, 2000) has been extended beyond serial short-term memory to tasks that require semantic processing (Marsh et al., 2008, 2009; Sörqvist, et al., 2012) wherein it is proposed automatic semantic processing of task-irrelevant sound interferes with deliberate semantic processing of focal task material (Marsh et al., 2008, 2009).

The interference-by-process view of disruption to serial short-term memory by to-be-ignored sound correctly accounts for why the post-categorical content of to-be-ignored sound is not influential in distraction of serial recall performance. Neither the meaningfulness of to-be-ignored sound (either whether or not it comprises words in a language a participant understands [Jones et al., 1990], or nonwords as compared with words [Salamé & Baddeley, 1982], nor between-sequence semantic similarity [similarity in meaning between to-be-ignored sound and visual memoranda; Buchner et al., 1996]) impairs serial recall. Serial recall tasks by design minimises participants' use of syntactical or lexical semantic processing that might aid ordered recall. In this way participants co-opt vocal-motor processing and paralinguistic skills to graft transitional probabilities onto items that do not have such properties. Motor-processing using serial order is open to interference by other serially-ordered material that may at some level fit the action-parameters. For example, if similar sounds, although containing spectral differences, are unified into a single stream, the order of such sounds are obligatory pre-attentively itemized. These involuntary cues can interfere with similar, deliberate, serial processing of to-be-remembered items in the vocal-motor plan (Hughes & Marsh, 2017).

While the interference-by-process account holds that the ISE is produced through processing acoustic features, it is possible that the post-categorical nature of sound, for example, a familiar melody accompanied by lyrics compared to an unfamiliar melody accompanied by lyrics, could produce greater disruption by drawing on the same neural substrates responsible for serial order retention. Serial recall arguably requires involvement of the supplementary motor area (SMA) for motor-planning and thus serial rehearsal of a to-be-remembered visual sequence (Lima, Krishnan, & Scott, 2016). Indeed, a number of studies have shown that overhearing a familiar melody also activates the SMA (Halpern & Zattore, 1999; Koelsch et al., 2009). Thus, a familiar melody could monopolize motor-planning systems that serial recall draws upon resulting in subsequent impaired performance. However, there is also an argument that predictive coding and schema-driven processing might happen for familiar melody resulting in stronger order cues and thus more interference-by-process. Although any impact on the SMA of a familiar, as compared to an unfamiliar melody, and the effect on sub-vocal motor-output organization, is not clear, added disruption from a melody that is familiar could indicate an effect of interference-by-process due to a stronger representation of order cues from overlearned memory for the melody.

5.2.2 Auditory distraction type II: Attentional capture.

According to the duplex account (Hughes, 2014), the second form of auditory distraction is attributable to attentional capture, and this form of distraction can be differentiated from the interference-by-process form (as indexed by the changing-state effect) in a number of ways (Hughes et al., 2007; Marsh et al., 2019). For example, in addition to being task-sensitive, the changing-state effect, contrary to attentional capture effects, is inviolable to top-down control. Increased task difficulty, or forewarning, does not eliminate the changing-state effect but reduces or eliminates effects that are produced by attentional capture (Hughes et al., 2013). Furthermore, unlike auditory distraction effects produced by interference-by-process, attentional capture effects can occur regardless of the type of task-processing adopted for the focal task and manifests when attention becomes momentarily disengaged from the task at hand (Hughes et al., 2005, 2007; Parmentier & Andrés, 2010; Parmentier et al., 2008).

Two types of attentional capture can be differentiated. First, a sound may capture attention due to its specific content. Thus, *specific attentional capture* occurs from one's own name (as compared to a yoked control name; Conway, Cowan, & Bunting, 2001; Moray, 1959; Röer et al., 2013; Wood & Cowan, 1995), taboo words (Röer et al., 2013), emotionally valent words (Buchner, Mehl, Rothermund, & Wentura, 2006; Marsh et al., 2018), sounds of interest to a given individual such as an overheard telephone conversation (Marsh et al., 2018), or food-related words for a hungry person (Parmentier, Pacheco-Unguetti, & Valero, 2018). Second, sound may capture attention because of the context within which it is heard as opposed to it possessing anything in and of itself that is attention-capturing. Examples of this *aspecific attentional capture* include potency of the letter sound "A" to capture attention following presentation of a sequence of different letter sounds "CCC**A**CC". The notion being that it violates expectation that another "C" will be presented. Any disruption, therefore, has nothing to do with the particular properties of the letter "A", since it would be "C" that would capture attention in a sequence "AAA**C**AA". This disruption, caused by such an acoustic change, has been termed the auditory deviation effect (Hughes et al, 2005; Vachon et al., 2012).

It should be mentioned here that while the duplex account asserts that attentional capture and the changing-state effect are underpinned by two different mechanisms, the unitary account (Bell et al., 2012; Cowan, 1995) proposes that both are explicable through the attentional capture concept. To explain the changing-state effect, the unitary account proposes that each acoustic change within a sequence triggers an orienting response resulting in attention being captured away from the focal task (Sokolov, 1963; Elliot & Cowan, 2001). Thus, each item within the sequence "FQRLYS" triggers an attentional capture while only the letter "L" in the sequence "FFF**L**FF" captures attention.

While an in-depth discussion of the unitary account is beyond the scope of the current chapter (see Chapter II), it has been undermined by several lines of work. For example, the account fails to explain why an auditory deviation in a changing-state sequence, "FQRLYS" produces the same magnitude of disruption compared to an auditory deviation (e.g., a change of voice) in a steady-state sequence "FFF**L**FF" (Hughes et al., 2013). On the unitary account the deviation effect should be smaller in the context of a changing-state sequence wherein each item within the sequence repeatedly captures attention anyway. In addition, the unitary account cannot explain why the changing-state effect is not observed on the missing-item task when this task is susceptible to disruption via the auditory deviation effect (Hughes et al., 2007). Furthermore, the unitary account cannot explain a number of key findings (e.g., associated with presentation of foreknowledge, Working Memory Capacity and sensory/encoding load) that show differential involvement of cognitive control modulates the auditory deviation effect but not the changing-state effect (Hughes, 2014). Due to well-expounded problems with the unitary account, characterization of attentional capture in the following assumes that it is a qualitatively different effect from the changing-state effect.

Previous studies have shown aspecific attentional capture in relation to perception of familiar song. For example, in an oddball task, categorization of arrows as left-or right-facing was delayed by the presence of a pitch violation to a familiar song "Twinkle, twinkle, little star". This disruption occurred irrespective of whether the violation was against or within the expected direction of pitch contour (Nöstl, Marsh, & Sörqvist, 2012). Other work demonstrating a double dissociation in rule and memory for music (Miranda & Ullman, 2007) suggest both pitch in-key and out-of-key deviants to familiar melodies could be attention-related. To date, disruption attributable to aspecific attentional capture from violations to an expected pattern based upon long-term knowledge for melody has not been observed in short-term memory tasks, while disruption from manipulations of proverb-end-words—that require post-categorical problem for identification—do indeed disrupt short-term memory (Röer, Bell, Körner, & Buchner, 2019). In any case, additional disruption attributable to familiar melody with familiar or unfamiliar lyrics relative to other auditory conditions used in Study I-II cannot be attributable to aspecific attentional capture since no auditory deviations were included within auditory distracter sequences.

If disruption attributable to familiar melody with familiar/unfamiliar lyrics is indeed attributable to attentional capture, then it could cohere with the notion of specific attentional capture. Like the own-name effect (Conway et al., 2001; Röer et al., 2013), a familiar melody may hold personal significance to participants, and, akin to valent words (Buchner et al., 2006; Marsh et al., 2018), may induce affective responses that capture a participant's attention away from the focal task.

On the face of it, disruption identified in Study I-II appears consistent with the view that irrelevant familiar melody/lyrics compete for vocal-motor apparatus required for production of a familiar melody (Study I) and lyrics via associative-cuing (Study II). Arguably, this arises because written retrieval cues can be underspecified thereby activating other memories that compete for recall (Anderson, 2003). Such a selection problem being compounded by the presence, within to-be-ignored sound, of familiar, as compared to unfamiliar, melody/lyrics. This being so, then patterns of disruption observed should be dissociated by task demands. That is, any additional disruption observed for familiar melody with lyrics over other auditory conditions in Study I should not obtain in a serial recall task that does not necessitate reproduction of melody.

However, disruption produced by familiar melody/lyrics may be attributable to specific attentional capture driven by personal significance/affective response to the auditory stimulus (cf. Conway et al., 2001; Buchner et al., 2006; Marsh et al., 2018;). If so, then disruptive effects observed in Study I-II should also be observed irrespective of whether the focal task entails reproduction of a melody. To this end, Study III explored whether the same patterns of disruption from the auditory distracter conditions adopted for Study I-II occurred for a visual-verbal short-term memory serial recall task that, like melody recall (Study I) and lyrics recall (Study II) requires vocal-motor processing, but, unlike those tasks does not require access to, or production of, melody/lyrics.

At a general level, if results are to be consistent with the interference-by-process component of the duplex account, then it was hypothesized that the changing nature of acoustic properties of sound condition should produce disruption (Jones, 1993; Jones & Tremblay, 2000). Moreover, in this view, speech being more acoustically complex and variable than tones, should produce a greater magnitude of disruption of serial recall. Furthermore, the interference-by-process view supposes that unlike Study I-II, the combination of familiar melody with lyrics would not disrupt performance more than unfamiliar melody with lyrics. This is since serial recall of digits does not require retrieval of melody and thus there should be no clash of processes. Alternatively, on the (specific) attentional capture view of the duplex account, there should be no relationship between sound properties and those of task: The serial recall task should be as disrupted by the same properties of sound that were most disruptive to melody recall (Study I) and lyrics recall (Study II). Namely, familiar lyrics with melody accompaniment should produce greater disruption to serial recall performance. Notably this outcome would mean that patterns of distraction observed would fail to provide an insight into lexical/lyrical/melodic retrieval processes.

5.3 Experiment 3: Method

5.3.1 Participants.

Two hundred participants undertook this Experiment. Participants comprised 26 Men and 74 Women, aged 18-88 (M = 46.55, SD = 22.68), recruited from U3A groups, Soroptimist, Rotary, choirs, orchestras, universities, colleges, nurseries, local community groups, and personal contacts, following email requests to their activities co-ordinators, poster display, and visits to meetings. Ethical approval, consent, and reimbursement procedure detailed in Chapter III, 3.2.1.

Cognition.

In keeping with experiments in these empirical series participants aged over 50 years undertook the Addenbrooke's Cognitive Examination Revised (ACE-R) which includes mini-mental state examination (MMSE) information[38]. Scoring ranged between 82-98 (M = 93.17, SD = 4.02), with MMSE ranging between 25-30 (M = 28.94, SD = 1.10). This indicates positive cognitive functioning. A mixed ANOVA showed that for Sound Type Group 2, Familiar Lyrics and Speech, the Sound Condition × Cognition[39] interaction effect was non-significant, $F(44, 152) = 1.022$, $MSE = .007$, $p = .447$, $\eta_p^2 = .228$. However, for Sound Type Group 1, Melody and Unfamiliar Lyrics, the Sound Condition × Cognition interaction was significant, $F(52, 144) = 1.783$, $MSE = .007$, $p = .004$, $\eta_p^2 = .392$. Pairwise comparisons showed all sound pairings to be significant except for between Familiar Lyrics and Unfamiliar Lyrics Groups (MD = .001, SE = .021, p = .958; 95% CI [-.041, .043]). This suggests that when combined with unfamiliar lyrics familiarity of the melody was unimportant irrespective of cognitive ability[40]. A visual acuity test for participants aged 50+ confirmed all had acceptable vision (task font 16), and a test for hearing loss, conducted via an audiometer, recorded hearing levels >30dB at 0.5-4.0 kHz for all older participants[41].

---

[38] <88 cut off giving 0.94 sensitivity for dementia with <82 giving 0.84. MMSE >24 no cognitive impairment.
[39] Based on ACE-R data.
[40] See scatterplots for correlation between recall scores and cognition in Appendix D.
[41] Normal hearing established at 0-25 dB, some mild hearing loss at whisper level recorded at 20-40 dB.

Musical culture.

From a demographic response sheet completed prior to beginning Experiment 3, a 'Western' musical culture was indicated by 95% of participants, with 3% indicating Asian, 1% Oriental, and 1% dual Western/African cultures.

Musical training and participation.

A repeated measures ANOVA comparing recall scores of all participants with musical training and current musical participation (as indicated on their demographic response sheet) indicated that the between participants effects and all interactions of Sound Condition × Training, Sound Condition × Instrument, and Sound Condition × Choir were non-significant. In total 29% of participants were deemed to be musicians[42] with 48% indicating instrumental ability. Results confirmed no significant musical advantage or disadvantage for the current serial recall task.

5.3.2 Design.

Experiment 3 was conducted as a mixed 5 (Sound Condition) × 2 (Sound Type) design, the dependent variable being accuracy of digit recall in strict serial order. The within-participants variable of Sound Condition was classified into five levels for each of two between participants Sound Type Groups: Group 1, Melody and Unfamiliar Lyrics; Group 2, Familiar Lyrics and Speech. For Group 1 sound levels comprised quiet (no distracter), Melody-familiar, Melody-unfamiliar, Unfamiliar sung-lyrics-familiar melody, Unfamiliar sung-lyrics-unfamiliar melody distracters. For Group 2 sound levels comprised quiet (no distracter), Familiar-sung-lyrics-familiar melody, Familiar sung-lyrics-unfamiliar melody, Speech-familiar, Speech-unfamiliar distracters. An additional repeated measures ANOVA for each Group examined onset time taken to begin recall and a mixed ANOVA was computed to establish prior knowledge of the target and distracter melodies from a Recognition test. A shortened duration of each sound file from 32-seconds (used for Study I-II) to 10-seconds enabled two Sound Type Groups from Study I-II to be paired. (A 10-seconds retention having been adopted for previous studies e.g., Jones et al., 1993; Klatte, Lee, & Hellbrück, 2002.) All participants experienced the quiet condition, but then, to reduce demand characteristics, distracter conditions were split, participants alternately allocated to one Sound Type, see Table 5.1. In order to address repeated measure unsystematic variation, the order of experimental conditions was counterbalanced, a random order of digit patterns preferred to a fixed order, Table 5.2.

---

[42] 7+ years of musical training (Ericsson et al., 1993).

Table 5.1

Experimental Sound Conditions within Sound Type Groups1 and 2

Sound Type Group	*n*	distracter variable
1 and 2	200	quiet
1 Melody + Unfamiliar Lyrics	100	familiar melody (no lyrics) unfamiliar matched melody (no lyrics) unfamiliar sung-lyrics (familiar melody) unfamiliar sung-lyrics (unfamiliar matched melody)
2 Familiar Lyrics + Speech	100	familiar sung-lyrics (familiar melody) familiar matched sung-lyrics (unfamiliar melody) familiar spoken-lyrics (no melody) unfamiliar spoken-lyrics (no melody)

Table 5.2

Order of Sound Condition presentation for Sound Type Groups 1 and 2

Sound Type Group 1	order	condition	order	condition
1 quiet	**1**	1 2 5 3 4	**6**	4 3 5 2 1
2 Melody - f	**2**	2 3 1 4 5	**7**	5 4 1 3 2
3 Melody - u	**3**	3 4 2 5 1	**8**	1 5 2 4 3
4 Unfamiliar Lyrics - f	**4**	4 5 3 1 2	**9**	2 1 3 5 4
5 Unfamiliar Lyrics - u	**5**	5 1 4 2 3	**10**	3 2 4 1 5

Sound Type Group 2	order	condition	order	condition
1 quiet	**1**	1 2 5 3 4	**6**	4 3 5 2 1
2 Familiar Lyrics - f	**2**	2 3 1 4 5	**7**	5 4 1 3 2
3 Familiar Lyrics - u	**3**	3 4 2 5 1	**8**	1 5 2 4 3
4 Speech - f	**4**	4 5 3 1 2	**9**	2 1 3 5 4
5 Speech - u	**5**	5 1 4 2 3	**10**	3 2 4 1 5

Note. Group 1: Melody and Unfamiliar Lyrics, Group 2: Familiar Lyrics and Speech. f - familiar melody, u - unfamiliar matched melody.

The criteria for pairing Sound Type within the Groups was to balance familiarity. It was considered that a familiar melody and a familiar song (familiar melody with familiar lyrics) had the greater familiar components and so should be allocated to different Groups. Moreover, the Melody and Speech conditions, each comprising one song component, were also allocated to separate Groups.

5.3.3 Materials and apparatus.

Ten familiar distracters for each Sound Type, were taken from nursery rhymes or traditional songs, the melodies of which all had associated lyrics, and together with ten unfamiliar matched distracters were developed, delivered, and executed in identical fashion to Study I-II (see Chapter III, 3.2.3).

5.3.4 Procedure.

Participants were tested individually in a quiet room, in the presence of the researcher, having previously completed consent and demographic forms. Following four practice trials in quiet conditions participants were visually presented via a computer screen with series of eight single digits drawn from 1-8. They were instructed to memorise them in order of presentation in quiet conditions and whilst being distracted through headphones (Sennheiser HD201) by irrelevant melody, a cappella song, or speech (spoken song lyrics), both familiar and unfamiliar, which they were instructed to ignore. There was no repetition of a digit within each trial. Each participant received a set of 50 trials across five conditions (see Table 5.1).

To ensure standardisation each digit series required a click on visually presented words "Begin Trial" to move onto the next series. Instructions were consistent for each movement through the program and prior to beginning each task a visual **+** was shown in the centre of the screen to focus the eye and alert participants as to the position of to-be-remembered information. Digits were presented consecutively at a rate of 1 per second (800-millisecond on, 200-millisecond off) in 72-point equidistant black Monaco font on a white background. From a viewing distance of 45 centimeters each number subtended a vertical visual angle of 1.49° and a horizontal angle of 0.92°.

For distraction, ten familiar nursery rhymes and ten novel nursery rhymes, matched for musical complexity[43], were presented through headphones via the computer, each .WAV file with a 10-seconds duration. Onset of to-be-ignored auditory distracter conditions was simultaneous with presentation of each visual digit pattern. Participants were instructed via the computer screen to ignore sounds sometimes heard through their headphones as they were irrelevant to requirements necessary for task completion.

Following a 2-seconds retention interval at the end of each 1000-millisecond trial, participants were presented on screen with a circular array of all 1-8 digits and required to indicate their correct sequence by clicking on each digit in order of presentation using the left mouse button. They were told to use **'?'** in the centre of the

---

[43] Comprising identical musical elements to those in Study I-II as detailed in Chapters III-IV.

circle of digits if they could not recall a digit in a specific position. The pattern of digits within the circle changed for each trial to eliminate practice order effect and address variation in onset time of to-be-remembered digits. There was no time-limit on recall, the next series beginning when all response boxes were completed by digit or **'?'**. A 30-second rest between experimental conditions was permitted if needed and this proposal was also indicated via the computer screen.

To control for prior knowledge of the ten familiar nursery rhyme distracter songs and ten unfamiliar distracter songs (to establish a baseline) participants were played the 20 melodies and 20 song lyrics—as speech—once their five conditions had been completed. Participants were requested to press the "Y" key if familiar and the "N" key if not familiar. The order being reversed for alternate participants to combat possible cueing of melody via lyrics or vice-versa. They were then instructed to indicate confidence in their decision by clicking on a point across a continuum ranging from "Not confident" to "Totally confident" the slide positioned at the centre of the continuum at the start of each trial (this Likert style format having a precedent for determining preferred or familiar music e.g., McCorkell, 2012; Perham & Currie, 2014). Each total task took approximately 45 minutes.

5.3.5 Assessment.

Study III aimed to identify the effect on accuracy of serial-digit recall by the presence or absence of a familiar or unfamiliar sound distracter by comparing means for each of nine levels of Sound variable. Accuracy of serial-digit recall was recorded by the E-Prime program using a scale ranging from zero, for no digits, to eight for all digits correctly recalled. Raw data scores were assessed on strict serial recall criterion: for a correct score, an item's recall serial position and presentation serial position needed to be identical. Data from each .edat file was then pasted into an Excel spreadsheet containing a formula for calculating the serial recall accuracy mean, data from each participant having been collapsed across each of their five conditions. A separate formula was entered to determine prior knowledge of the 20 song melodies/lyrics and participants' confidence in their decision.

5.4 Results

Following a one-way repeated measures ANOVA[44], serial-digit recall scores were found to be more accurate in quiet conditions (see Figure 9). An unexpected inconsistency in recall score accuracy in the quiet condition between Group 1 and

---

[44] Huynh-Feldt reported following sphericity violation ε =.864

Group 2 created a slight baseline slip. (Group 2 appear to be more able.) However, for both Groups, scores in quiet conditions surpassed those from any distracter conditions to a significant effect (Figure 9, Panel A and Panel B). From the eight distracter conditions, scores from the Melody only Sound Type, irrespective of familiarity, were shown to be the highest, but were closely followed by the Speech Sound Type. However, for both Melody and Speech a familiar sound was more detrimental to serial recall than an unfamiliar sound.

Recall scores decreased in Familiar Lyrics Sound Type combination conditions, a further decrement to scores noted during Unfamiliar Lyrics Sound Type. Interestingly, there was a pattern across lyrics conditions. An unfamiliar melody impaired recall more than a familiar melody when heard in combination with lyrics irrespective of the familiarity of those lyrics. However, when Melody or Speech were heard independently, the familiarity effect was reversed. Unfamiliar Melody and unfamiliar Speech were less detrimental to recall accuracy than familiar Melody or familiar Speech. However, unfamiliar sung-lyrics to an unfamiliar melody, decreased performance the most.

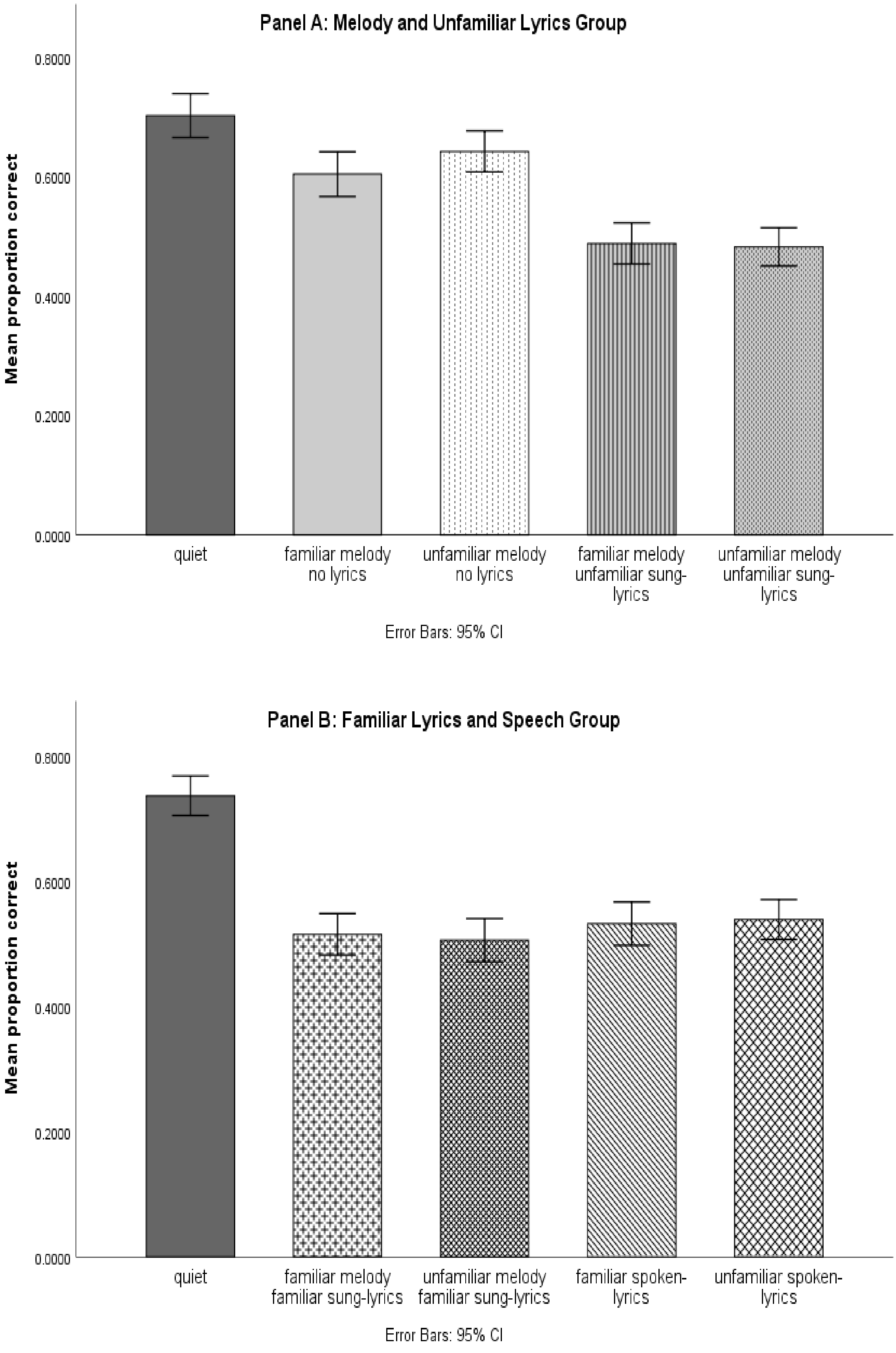


Figure 9. Panel A: Mean serial recall performance across Sound Conditions Group 1. Panel B: Mean serial recall performance across Sound Conditions Group 2.

For Group 1 (Melody and Unfamiliar Lyrics) a one-way repeated measures ANOVA showed that there was a significant main effect of Sound Condition against quiet[45], $F(3.356, 332.210) = 95.185$, $MSE = .012$, $p < .001$, $\eta_p^2 = .490$. A simple effects analysis[46] identified all Sound pairings to be significant at or lower than $p < .001$ except for the two sung Unfamiliar Lyrics conditions, to either a familiar or unfamiliar melody, familiar melody scoring highest (*MD* = .006, *SE* = .011, *p* = .602; 95% CI [-.016, .028]). However, a familiarity effect was identified for Melody (*MD* = -.038, *SE* = .011, *p* = .001; 95% CI [-.059, -.017]), whereby familiar melody alone was more distracting than unfamiliar melody. This suggests melody familiarity was unimportant when combined with unfamiliar lyrics.

A two-way ANOVA comparing familiarity with unfamiliarity showed a main effect of familiarity, $F(1, 99) = 4.281$, $MSE = .006$, $p = .041$, $\eta_p^2 = .041$. It would appear that the addition of lyrics to a familiar melody (that were not the lyrics normally associated with that melody) created increased impairment as opposed to when the familiar melody was heard without lyrics. However, there was also a main effect for unfamiliarity, $F(1, 99) = 129.641$, $MSE = .015$, $p < .001$, $\eta_p^2 = .567$, that suggests the presence of an unfamiliar lyric impaired performance as opposed to no lyric being present. A simple effects analysis for familiarity to decompose the interaction (familiar melody with no lyrics *versus* familiar melody with unfamiliar lyrics), showed a significant effect (*MD* = -.016, *SE* = .008, *p* = .041; 95% CI [-.032, -.001]). There was also a significant effect for unfamiliarity, no lyrics compared to unfamiliar lyrics (*MD* = .138, *SE* = .012, *p* < .001; 95% CI [.114, .162]). Furthermore, there was a significant familiarity × unfamiliarity interaction effect between melody familiarity and unfamiliarity (addition of unfamiliar lyrics), $F(1, 99) = 8.333$, $MSE = .006$, $p = .005$, $\eta_p^2 = .078$. A simple effects analysis showed a significant difference between familiarity and unfamiliarity (*MD* = .038, *SE* = .011, *p* = .001; 95% CI [.017, .059]). Therefore, the presence of lyrics (unfamiliar) impaired serial recall performance regardless of familiarity of the melody. Moreover, a familiar melody only resulted in poorer performance than an unfamiliar melody when there were no lyrics present, and not when lyrics were included (*MD* = .006, *SE* = .011, *p* = .602; 95% CI [-.016, .028]).

The significant main effect of Sound Condition was replicated for Group 2 (Familiar Lyrics and Speech), $F(4, 396) = 112.387$, $MSE = .008$, $p < .001$, $\eta_p^2 = .532$. A

[45] Huynh-Feldt reported following sphericity violation $\varepsilon = .839$
[46] LSD reported as it maintains statistical power. See Chapter III, 3.2.5.

simple effects analysis showed Sound effects across Sound Conditions was determined by sung *versus* spoken-lyrics Sound Type Groups. All Sound pairings were significant ($p < .05$) except for the Familiar Lyrics Sound Type (combined with a familiar or unfamiliar melody), that suggests when distracted by Familiar Lyrics melody familiarity was unimportant, see Table 5.3. No significant effect occurred between familiar and unfamiliar Speech suggesting equivalence in distractibility. Therefore, lyrics irrespective of delivery, spoken or sung, impaired serial recall to a similar extent irrespective of familiarity.

Table 5.3

Pairwise comparison across Sound Conditions showing non-significant results for Group 2

Familiar Lyrics	Sound Condition	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
	familiar melody *vs.* unfamiliar melody	.009	.013	.462	-.016	.035
	vs. familiar Speech	-.017	.012	.166	-.041	.007
Speech						
	familiar *vs.* unfamiliar	-.007	.012	.568	-.030	.017

A two-way ANOVA for Group 2 showed no main effect of familiarity regardless of whether lyrics were sung or spoken, $F(1, 99) = .026$, $MSE <.001$, $p = .871$, $\eta_p^2 < .001$. There was also no familiarity × unfamiliarity interaction, $F(1, 99) = .767$, $MSE = .008$, $p = .383$, $\eta_p^2 = .008$. However, there was a significant unfamiliarity effect, $F(1, 99) = 9.283$, $MSE = .007$, $p = .003$, $\eta_p^2 = .086$. Pairwise analysis showed this to be driven by lyrics that impaired performance when sung in the presence of an unfamiliar melody (*MD* = -.025, *SE* = .008, $p = .003$; 95% CI [-.041, -.009]). A simple effects analysis comparing unfamiliarity (familiar sung-lyrics as compared to unfamiliar spoken-lyrics) confirmed the significant effect between familiar sung-lyrics (with an unfamiliar melody) and unfamiliar spoken-lyrics. The greater incongruity the greater impairment to focal task (*MD* = -.033, *SE* = .013, $p = .010$; 95% CI [-.058, -.008]).

This experimental design necessitated lyrics familiarity to be between-participants, unfamiliar lyrics present in Group 1 (Melody and Unfamiliar Lyrics Group), familiar lyrics present in Group 2 (Familiar Lyrics and Speech Group). Therefore, a further two-way mixed ANOVA was computed to ascertain if there was any interaction between melody familiarity and lyrics familiarity. There was none. There was no significant effect for familiarity of the melody, *F*(1, 198) = .811, *MSE* = .007, *p* = .369, η_p^2 = .004, and no familiar melody × lyrics interaction, *F*(1, 198) = .047, MSE = .007, *p* = .829, η_p^2 < .001. The between-participant main effect of lyrics familiarity was also not significant, *F*(1, 198) = 1.344, *MSE* = .050, *p* = .248, η_p^2 = .007. It would appear that impairment to serial-digit recall occurred irrespective of melody or lyrics familiarity, when combined within song.

5.4.1 Recognition and confidence.

Participants indicated their familiarity of the nursery rhyme melody distracters (maximum 10 from each condition), and speech distracters (maximum 10 from each condition). For melody distracter Recognition, total scores for the Melody and Unfamiliar Lyrics Group indicated a mean of 8.6/10 (*SD* = 1.75; *SE* = .175) correct identification of familiar melodies (Hits), and 8.54/10 (*SD* = 1.73; *SE* = 1.72) correct rejections (CR) of unfamiliar matched melodies. This yielded a hit rate of .86 (*SD* = .175; *SE* = .017) and a false alarm rate of .146 (*SD* = .173; *SE* = .017). For the Familiar Lyrics and Speech Group, total scores for melody distracter Recognition demonstrated a mean of 8.5 (*SD* = 1.85; *SE* = .185) correct identification of familiar melodies (Hits), and 8.73 (*SD* = 1.483; *SE* = .148) correct rejections (CR) of unfamiliar matched melodies. The ensuing hit rate was .85 (*SD* = .185; *SE* = .019), with a false alarm rate of .127 (*SD* = .148; *SE* = .015).

For recognition of spoken-lyrics distracters, total scores for the Melody and Unfamiliar Lyrics Group showed a mean of 8.25/10 (*SD* = .198; *SE* = .198) correct identification of familiar spoken-lyrics distracters (Hits), and 8.66/10 (*SD* = 1.56; *SE* = .156) correct rejection (CR) of unfamiliar spoken-lyrics distracters. This led to a hit rate of .825 (*SD* = 1.98; *SE* = .020), and a false alarm rate of .134 (*SD* = .158; *SE* = .016). For Familiar Lyrics and Speech Group, total scores for recognition of familiar spoken-lyrics distracters was 8.72/10 (*SD* = 1.46; *SE* = .146) for correct identification of familiar spoken-lyrics distracter (Hits), and 8.77 (*SD* = 1.53; *SE* = .153) for correct rejections (CR) of unfamiliar spoken-lyrics distracters. The resultant hit rate was .872 (*SD* = .198; *SE* = .020), with a false alarm rate of .134 (*SD* = .158; *SE* = .016). A 2 (Stimulus Type: Melody *vs.* Speech) × 2 (Rate: Hit Rate *vs.* False Alarm) × 2 (Group: Melody and Unfamiliar Lyrics *vs.* Familiar Lyrics and Speech) revealed no main effect of Stimulus Type, *F*(1, 198) = .64, *MSE* = .017, *p* = .42, η_p^2 = .003, or of Group, *F*(1, .198) = .019, *MSE* = .035, *p* = .89, η_p^2 = .000. There was a main effect of Rate, *F*(1,

.198) = 2584.66, *MSE* = .04, $p < .001$, $\eta_p^2 = .929$, owing to finding that hit-rate was significantly greater than false alarm rate. Two-way interactions between Stimulus Type and Group ($F = 3.19$, $p = .076$), Rate and Group ($F = 1.386$, $p = .24$), Stimulus Type and Rate ($F = .007$, $p = .943$), and the three-way interaction between Stimulus Type, Rate, and Group ($F = 1.44$, $p = .232$) were not significant.

D'prime and criterion shift (c) were computed for recognition scores. However, this led to a lot of data loss since several participants did not commit a miss response (27 in the Melody and Familiar Lyrics Group, and 17 in the Familiar Lyrics and Speech Group for the melody condition, and for the spoken-lyrics condition, 23 in the Melody and Unfamiliar Lyrics Group and 25 in the Familiar Lyrics and Speech Group). For 11 participants in the Melody and Unfamiliar Lyrics Group and for 23 in the Familiar Lyrics and Speech Group a miss response was not committed in both melody and speech conditions[47]. For those participants d'prime and criterion shift measures could not be calculated. In the melody condition, d'prime scores were 1.45 (*SD* = .754; *SE* = .124) for Melody and Unfamiliar Lyrics Group, and 1.65 (*SD* = .559; *SE* = .095) for Familiar Lyrics and Speech Group. In the spoken-lyrics condition scoring was 1.54 (*SD* = .959; *SE* = .160) for Melody and Unfamiliar Lyrics Group, and 1.51 (*SD* = .661; *SE* = .117) for Familiar Lyrics and Speech Group. Too few data remained to perform a 2 (Stimulus Type: melody *vs.* spoken-lyrics) × 2 (Group: Melody and Unfamiliar Lyrics *vs.* Familiar Lyrics and Speech) analysis of variance on the d'prime scores due to the fact that d'prime scores could not always be computed at each level of Stimulus Type. A univariate analysis on the d'prime scores for melody showed no difference between the Melody and Unfamiliar Lyrics Sound Type Group ($n = 37$) and the Familiar Lyrics and Speech Group ($n = 35$), $F(1, 70) = 1.576$, *MSE* = .444, $p = .214$, $\eta_p^2 = .022$. Similarly, a univariate analysis on the d'prime scores for the spoken-lyrics stimuli showed no difference between Melody and Unfamiliar Lyrics Group ($n = 36$), and Familiar Lyrics and Speech Group ($n = 32$), $F(1, 66) = .027$, *MSE* = .693, $p = .869$, $\eta_p^2 = .000$, suggesting participants were able to discriminate familiar from unfamiliar melody/lyrics regardless of their Sound Type Group, see Figure 10, Panel A and Panel B.

[47] Resulting from 100% Hit/CR scores

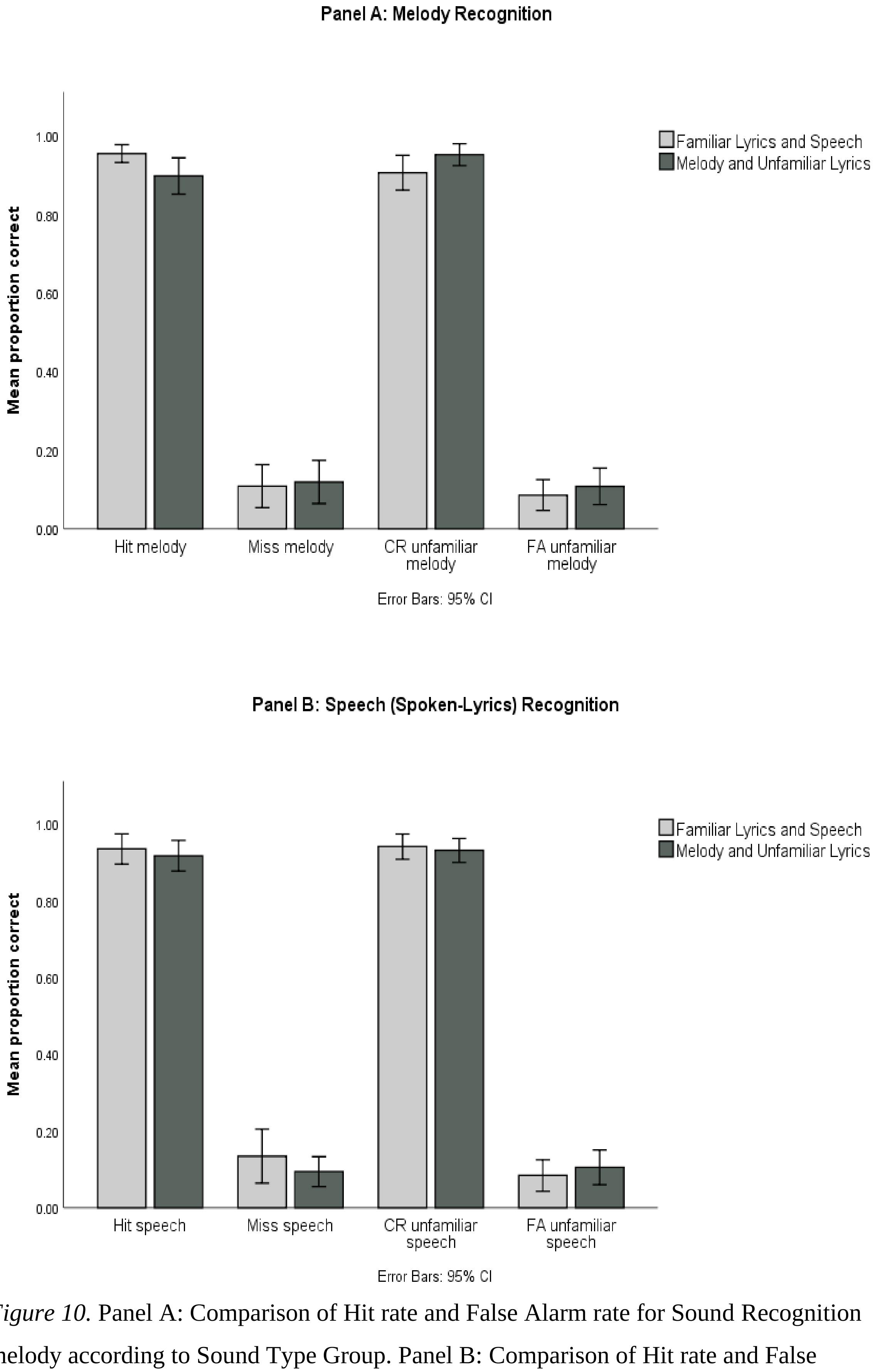


Figure 10. Panel A: Comparison of Hit rate and False Alarm rate for Sound Recognition melody according to Sound Type Group. Panel B: Comparison of Hit rate and False Alarm rate for Sound Recognition speech (spoken-lyrics) according to Sound Type Group.

For the melody condition c scores were .142 (SD = .347; SE = .057) for Melody and Unfamiliar Lyrics Group, and -.072 (SD = .326: SE = .055) for Familiar Lyrics and Speech Group. For the spoken-lyrics condition c scores for Melody and Unfamiliar Lyrics Group were -.001 (SD = .290; SE = .048), and for Familiar Lyrics and Speech Group .029 (SD = .332; SE = .059). A univariate analysis on c scores for melody showed a significant difference between Melody and Unfamiliar Lyrics (n = 37) and Familiar Lyrics and Speech (n = 35) Groups, $F(1, 70) = 7.255$, $MSE = .113$, $p = .009$, $\eta_p^2 = .094$, in that response bias was greater in the Melody and Unfamiliar Lyrics Group wherein a conservative response criterion was adopted while a neutral response criterion was adopted within the Familiar Lyrics and Speech Group. In contrast, a univariate analysis on c scores for spoken-lyrics showed no difference between the Melody and Unfamiliar Lyrics (n = 36) and the Familiar Lyrics and Speech (n = 32) Groups, $F(1, 66) = .158$, $MSE = .096$, $p = .692$, $\eta_p^2 = .002$.

For confidence scores, participants who made at least one response in each of the four categories for melody and spoken-lyrics stimuli were not necessarily the same participants, therefore, two 4 (Decision Type) × 2 (Group) mixed ANOVA's were computed for melody and spoken-lyrics. For the melody stimuli, there was a significant effect for Decision Type[48], $F(2.176,152.340) = 20.730$, $MSE = 36360.159$, $p < .001$, $\eta_p^2 = .228$. The between participants effect for Group was non-significant, $F(1.70) = .404$, $MSE = 61590.826$, $p = .527$, $\eta_p^2 = .006$, and there was a non-significant interaction between Decision Type and Group, $F(2.176,152.340) = .288$, $MSE = 36360.159$, $p = .769$, $\eta_p^2 = .004$. Pairwise comparisons for Decision Type showed all pairings to be significant except for Miss *versus* CR and Miss *versus* FA, see Table 5.4. For the spoken-lyrics stimuli, there was a significant effect for Decision Type[49], $F(2.255,148.831) = 20.242$, $MSE = 24559.071$, $p < .001$, $\eta_p^2 = .235$. The between participants effect for Group was non-significant, $F(1.66) = .675$, $MSE = 54082.569$, $p = .414$, $\eta_p^2 = .010$, and there was a non-significant interaction between Decision Type and Group, $F(2.255,148.831) = .684$, $MSE = 24559.071$, $p = .523$, $\eta_p^2 = .010$. Pairwise comparisons for Decision Type only showed significance against Hit, see Table 5.4 suggesting increased confidence for matched spoken-lyrics than for melody.

[48] Huynh-Feldt reported following sphericity violation $\varepsilon = .725$
[49] Huynh-Feldt reported following sphericity violation $\varepsilon = .752$

Table 5.4

Pairwise confidence comparisons for melody and speech (spoken-lyrics) stimuli

Stimuli	Decision Type	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Melody	Hit *vs*. Miss	166.014	32.522	< .001	101.151	230.877
	Hit *vs*. CR	136.288	27.881	< .001	80.681	191.894
	Hit *vs*. FA	198.251	21.900	< .001	154.573	241.930
	Miss *vs*. CR	-29.727	18.422	.111	-66.467	7.014
	Miss *vs*. FA	32.237	31.623	.312	-30.832	95.307
	CR *vs*. FA	61.964	27.305	.026	7.507	116.421
Speech	Hit *vs*. Miss	162.688	24.075	< .001	114.620	210.756
	Hit *vs*. CR	137.522	21.730	< .001	94.136	180.908
	Hit *vs*. FA	140.160	19.603	< .001	101.022	179.298
	Miss *vs*. CR	-25.166	16.825	.139	-58.758	8.425
	Miss *vs*. FA	-22.528	29.235	.444	-80.898	35.843
	CR *vs*. FA	2.638	26.366	.921	-50.003	55.280

Note. CR - correct rejection; FA - false alarm.

5.4.2 Response times.

In addition to recall scores, response time to begin each digit series was measured. Response times showed a relative consistency between conditions see Table 5.5, the exception being longer response time in lyric conditions, especially for the Familiar Lyrics Group in the familiar melody Sound Condition.

Table 5.5

Response time mean scores and standard deviation for Sound Type Groups

Group 1: Melody + Unfamiliar Lyrics	Mean *(SD)* *millisecond*	Group 2: Familiar Lyrics + Speech	Mean *(SD)* *millisecond*
quiet	1056.21 *(433.22)*	quiet	1059.31 *(389.57)*
Melody - familiar	1073.00 *(403.01)*	Familiar Lyrics - familiar melody	1177.52 *(510.28)*
Melody - unfamiliar	1071.41 *(372.30)*	Familiar Lyrics - unfamiliar melody	1146.53 *(481.03)*
Unfamiliar Lyrics - familiar melody	1064.39 *(374.06)*	Speech - familiar	1133.97 *(482.61)*
Unfamiliar Lyric - unfamiliar melody	1103.95 (*468.18)*	Speech - unfamiliar	1153.25 *(452.88)*

A mixed ANOVA showed that although there was no significant effect of Sound Condition for the Melody and Unfamiliar Lyrics Sound Type Group[50], $F(4, 339.819) = .913$, $MSE = 41896.075$, $p = .445$, $\eta_p^2 = .009$), there was a significant effect for the Familiar Lyrics and Speech Group[51], $F(3.600, 356.441) = 4.249$, $MSE = 52298.89$, $p = .003$, $\eta_p^2 = .041$). Pairwise comparisons showed that this was driven by quiet as compared to all distracter present conditions, see Table 5.6. No other distracter Sound pairings were significant, see Table 5.7. These results clearly indicate that time required to begin recall of target items in correct serial order was impaired to the greatest extent, in comparison to quiet, when items were presented in the presence of lyrics, sung or spoken. Interestingly, however, for Group 1, when unfamiliar lyrics were combined with familiar melody, recall time was not significantly impaired compared to quiet (*MD* = -8.189, *SE* = 28.792, *p* = .777; 95% CI [-.65.318, 48.941]), nor when combined with unfamiliar melody (*MD* = -47.743, *SE* = 30.148, *p* = .116; 95% CI [-107.564, 12.078]).

Table 5.6

Pairwise comparison across Sound Conditions showing significant results against quiet for Group 2

Sound Conditions	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
quiet *vs.* Familiar Lyrics - familiar melody	-118.218	36.860	.002	-191.355	-45.080
Familiar Lyrics - unfamiliar melody	-87.222	34.277	.012	-155.236	-19.208
Speech - familiar	-74.659	33.655	.029	-141.437	-7.881
Speech - unfamiliar	-93.945	30.572	.003	-154.607	-33.284

[50] Huynh-Feldt reported following sphericity violation ε = .858
[51] Huynh-Feldt reported following sphericity violation ε =.900

Table 5.7

Pairwise comparison across Sound Conditions showing non-significant results for Group 2

Sound Conditions	*MD*	*SE*	*sig*	Lower Bound	Upper Bound
Familiar Lyrics - familiar melody *vs.* Familiar Lyrics - unfamiliar melody	30.996	25.975	.236	-20.544	82.535
Familiar Lyrics - familiar melody *vs.* Speech - familiar	43.559	33.416	.195	-22.747	109.864
Familiar Lyrics - familiar melody *vs.* Speech - unfamiliar	24.273	28.961	.404	-33.193	81.738
Familiar Lyrics - unfamiliar melody *vs.* Speech - familiar	12.563	25.644	.625	-38.320	63.446
Familiar Lyrics - unfamiliar melody *vs.* Speech - unfamiliar	-6.723	25.907	.796	-58.129	44.683
Speech - familiar *vs.* Speech - unfamiliar	-19.286	29.243	.511	-77.310	38.738

5.5 Discussion

For Study III the main aim was to identify if distraction effects observed in Study I (melody recall) and Study II (lyrics recall) could be replicated in a visual-verbal serial recall task known to be vulnerable to disruption by task irrelevant sound (Jones et al, 1992; Jones & Tremblay, 2000; Salamé & Baddeley, 1989). Study III, therefore, would provide an opportunity to identify if results from Study I-II were task-process-specific, only achieved when participants were required to recall and perform components of song, or if they were task-process-independent, achieved irrespective of task requirement.

All auditory conditions were identical to those for Study I-II. However, there were three main modifications to procedure. First, for Study III all participants experienced five, as compared to the previous three, irrelevant Sound Conditions; second, each sound stimuli was reduced from 32 to 10-second to correspond to focal task duration requirement; and third, irrelevant sound occurred during short-term to-be-remembered item presentation, rather than during long-term memory retrieval performance.

5.5.1 Comparison of overall results with previous findings.

Overall, as expected, serial-digit recall accuracy was higher in quiet conditions as compared to all Sound Conditions, with or without lyrics, familiar or unfamiliar. This is in line with previous literature using music distraction in serial recall (Alley & Greene, 2008; Baddeley, 1986; Jones & Macken, 1993; McCorkell, 2012; Pring & Walker, 1994; Salamé & Baddeley, 1989). Single distracter conditions caused less impairment to focal task recall than combination conditions. In other words, participants were impaired less by melody without lyrics, unfamiliar or familiar, or by spoken-lyrics without melody, familiar or unfamiliar, than when melody and lyrics were heard in combination as song, familiar or unfamiliar. These findings, in part, support Iwanaga and Ito (2002), Nittono (1997), and Salamé and Baddeley (1989), who identified melody and lyrics combined within song to be more distracting than irrelevant melody, or irrelevant speech (Alley & Greene, 2008). In addition, results from this current study showed a distracter familiarity effect in the single Melody sound condition, whereby an unfamiliar sound was less distracting than a familiar sound. However, if this distracter familiarity effect is due to interference-by-process or attentional diversion was not resolved. This result adds to the familiarity debate.

5.5.2 Interference-by-content *versus* interference-by-process.

Serial order representations of a familiar melody distracter known to participants, may be activated to a greater degree than an unfamiliar distracter due to their embedded learning, thereby creating interference-by-process on serial tasks. For Study III, the melody familiarity effect, whereby a familiar melody was significantly more distracting than an unfamiliar melody, concurs with this view. However, this was not replicated when melody was combined with lyrics. Results from this present study showed greater impairment to serial-digit recall from irrelevant melody combined with sung-lyrics, arguably containing increased changing state shown to be particularly disruptive to serial recall (Jones, 1994). Therefore, the addition of familiar or unfamiliar lyrics suggests disruption from lyrics overrode any disruptive impact of familiar melody.

The idea that familiarity with an unfolding auditory distracter may reduce impairment on focal task performance (Röer et al., 2015) was not totally borne out by results from this current study. In particular, serial-digit recall accuracy in the familiar melody condition was significantly reduced as compared with the unfamiliar melody condition. This did not occur when melody was combined with lyrics. Although speculative, it is possible that familiar melodies have personal-significance or meaning for participants which captured their attention (like their own name, Röer et al., 2013).

Emotions generated by music genres/elements, such as familiar *versus* unfamiliar (Cassidy & McDonald, 2009; Feng & Bidelman, 2015), liked (Perham & Vizard, 2011), preference (Perham & Sykora, 2012), arousal (Schellenberg, 2005; Thompson et al, 2001), and stimulating *versus* sedative (Schlittmeier et al., 2008) have been shown to influence participant's ability to accurately complete tasks with concurrent irrelevant background music (but see Gustavson, Hanneken, Moldysz, & Simon, 2013). However, this idea cannot account for the melody familiarity effect's inability to be replicated when melodies were combined with their associated familiar lyrics. Distracter conditions across Study III were consistent. All stimuli were identical nursery rhymes comprising the same tonality, tempo, dynamic, and instrumentation, so they were unlikely to create diverse emotional impact or to influence condition group outputs. Mechanisms contributing to the melody familiarity effect were, therefore, unresolved.

Study III required rehearsal of non-musical items, digit sequences visually presented, and recall in order of presentation. This act was unrelated to concurrent presence of irrelevant auditory conditions. The idea that impairment results from requirement to remember and simultaneously ignore similar items that require similar seriation processes—interference-by-content—(Neath, 2000; Salamé & Baddeley, 1982) is not supported from these current results. However, other accounts propose disruption is influenced by acoustic variation (Jones & Macken, 1993) and auditory processing is reliant on segmentation (Macken et al., 2009) and streaming (Bregman, 2001). Irrespective of the nature of the sound, serial-digit recall in this current study was disrupted. Rather, findings support the idea that competition between obligatory and deliberate ordering of sequence patterns is paramount in causing impairment to focal task performance—interference-by-process—(Jones & Tremblay, 2000; Marsh et al., 2009).

In sum, Study III explored propensity for familiar auditory sequences to produce greater disruption of short-term memory than unfamiliar sequences, but results showed that performance was disrupted more by familiar distracters only when sequences comprised an instrumental melody. To-be-recalled and to-be-ignored items for this current study bore no similarity but the task did require rehearsal of order information concurrent with perceptual organization of disparate irrelevant stimuli. Results lie in favour with interference-by-process/perceptual gestural accounts (Hughes & Jones, 2005; Hughes et al., 2009).

5.5.3 Perceptual-gestural account.

According to the perceptual-gestural account the “perceptual-organization”, or order, of irrelevant sound is of greater importance than its content. The perceptual component of this account applies mainly to the auditorily presented sequence of items (Hughes et al., 2009). Alternatively, the gestural component applies equally to sequences of auditory and visual origin and is supported by studies of the irrelevant sound effect (Hughes & Jones, 2005) and the linguistic familiarity effect (Woodward et al., 2008) within the domain of short-term memory.

For Study III, seriation was necessary on two levels; first, to maintain order of item presentation during rehearsal, and second, auditory conditions involved obligatory seriation of the unfolding irrelevant melody or lyrics contour. Requirement to convert visual-verbal sequences into articulatory form, thereby exploiting the speech-planning mechanism, posits accuracy reliance on sub-vocal rehearsal to organise and maintain the required sequence rather than the offsetting of any time-dependent decay in a purported phonological store (Baddeley, et al., 1975). As involuntary processing of irrelevant sound also entailed seriation processing it could be argued that this conflict of order process was inherent in impairment to accuracy in the current study of serial-digit recall.

Study III results further illustrate how the specificity of task processes dictate the disruption produced by different properties of the task-irrelevant sound since the patterns of interference differed when recall in serial order was required. Crucially, in Study III, the combination of melody and lyrics with different levels of familiarity, unlike Study I-II, did not modulate the distraction of serial recall which coheres with the argument that the distracters assume disruptive power when they conflict with the focal task process. This suggests that the disruption arising from familiar melody/unfamiliar melody and familiar/unfamiliar lyrics to melody retrieval (Study I, humming) and lyric retrieval (Study II, speaking) was not simply the result of attentional capture. However, familiar melody impaired serial recall more than unfamiliar melody and the basis of this effect requires further investigation. Globally, these results fit within the framework of interference-by-process and perceptual-gestural account (e.g., Hughes et al., 2009, 2011, 2016) and undermine the attentional capture account (e.g., Cowan, 1995, Neath, 2000).

5.5.4 Attentional capture.

The unitary account posits that the changing-state effect results from attentional capture (Cowan, 1995, Neath, 2000). However, it is difficult to envisage how this version of attentional capture could support the finding from Study III, that serial recall accuracy demonstrated a significant decrease in the presence of a familiar against unfamiliar distracter within the context of the melody only condition, since both conditions should have similar levels of changing-state. Alternatively, it is possible that familiarity of the melody itself was capable of diverting attention away from the focal task, a specific form of attentional capture. However, the results from the sung-lyrics and spoken-lyrics conditions Study III contradict this theoretical viewpoint. Here, in all four sung-lyrics conditions an irrelevant familiar melody was not significantly detrimental to task performance than an irrelevant unfamiliar melody. Moreover, spoken-lyrics that were familiar did not impair recall accuracy any more than those that were unfamiliar. Therefore, familiarity per se was not driving attentional capture. Barring the distracter familiarity effect in the context of the melody only condition, these findings are more easily explicable within the interference-by-process account.

5.6 Task process independent or specific?

Findings from this current study show that irrespective of familiarity, melody and lyrics, combined within song as sung-lyrics, were not significantly more disruptive to serial-digit recall than spoken-lyrics. Moreover, familiarity of melody was unimportant when in combination with lyrics. Although at first glance it would appear that the significant pattern of results from Study I-II were not replicated for Study III, there were some similarities that warrant closer inspection.

Greater disruption to recall performance across the three tasks, from long-term or short-term memory, resulted from song conditions (Familiar sung-lyrics, and Unfamiliar sung-lyrics). The main difference being that when required to recall a component of song (Study I-II), familiarity of the irrelevant sung-lyrics was most detrimental (Familiar sung-lyrics scores lower than Unfamiliar sung-lyrics). This contrasted with the non-music serial-digit task whereby unfamiliarity of lyrics was most captivating causing greater disruption. (Unfamiliar sung-lyrics lowest scoring.) In addition, although both single conditions, Melody and Speech, were least disruptive to any recall tasks, Speech was more disruptive than Melody for semantic, lyrics recall (Study II) and serial-digit tasks (Study III). This would suggest irrelevant melody impeded a music task (Study I) to a greater extent than for semantic or non-music serial recall tasks.

This is in line with previous verbal (Le Compte et al., 1997; Salamé & Baddeley, 1989), and music literature (Pechmann & Mohr, 1992; Schendel & Palmer, 2007; Williamson et al., 2010) relating to a modality specific ISE when comparing tones and speech. For example, Williamson et al. (2010) showed irrelevant speech to be more detrimental to letter recognition than irrelevant tones, and irrelevant tones to be more detrimental than irrelevant speech to tone recognition. A plausible explanation for this modality specific ISE suggests that it occurs due to an overlap between physical characteristics of information streams (Schendel & Palmer, 2007). Taken together these results support a degree of independence in processing verbal and tonal material in short-term memory. Therefore, separation propounded by Salamé and Baddeley (1989) whereby digit recall accuracy decreased with greater irrelevant speech content in the irrelevant sound, is, in part, replicated here. The fact that familiar melody alone was less disruptive than any verbal condition points to some degree of separation at the semantic level for melody. Based on these results, the hypothesis that the auditory distraction effects are task-process specific/independent remains inconclusive.

5.7 Concluding summary and subsequent research

Familiar distracters produced more disruption of serial-digit task recall than unfamiliar distracters. However, familiar distracters were not more disruptive than unfamiliar distracters when combined with lyrics as song (Study I), suggesting that the disruptive from lyrics overrode the disruptive impact of familiar melody. While familiarity appears to make some difference to serial-digit recall this was not as pronounced as for Study I, melody recall. Study III, however, did provide some evidence for familiar distraction *per se* but only when distracters were heard separately (melody or speech alone), not in combination (as song). It appears from these results therefore, that the two effects are underpinned by different mechanisms. However, as debate continues as to whether the changing-state effect and attentional diversion are underpinned by the same mechanism (Körner et al., 2018), the identified familiarity effect from Study III needs to be teased out by comparing it in the context of a task, in the presence of the same irrelevant sound conditions, that does not require serial recall.

CHAPTER VI

EMPIRICAL STUDY IV: DISRUPTION OF SHORT-TERM MEMORY BY FAMILIAR MELODY: INTERFERENCE BY PROCESS OR ATTENTIONAL CAPTURE?

Abstract

Study III revealed that a familiar melody (without associated lyrics), produced more disruption of serial recall than a matched unfamiliar melody. There are potentially two explanations of this finding. First, on the perceptual-gestural view, processing of physical changes in a familiar sound may yield stronger order cues to conflict with serial-ordering processes appropriated to undertake the serial recall task (interference-by-process). Second, a familiar melody, as compared with an unfamiliar melody, may be more potent at diverting attention away from a visually-based focal task (attentional capture). As the distracter familiarity effect identified in Study III was observed only for the Melody Sound Type Group, this single Sound Condition was delivered during the current study in order to address whether the effect was indeed attributable to an interference-by-process. Study IV, therefore, investigated whether the disruptive effect of familiar melody over unfamiliar melody would be eliminated for a visually-based focal task that does not necessitate serial order processing. The rationale was that if the missing-item task, shown invulnerable to disruption via changing-acoustic properties of sound (Hughes et al., 2007), does not demand rehearsal of to-be-remembered items in order of presentation (Hughes et al., 2007; Jones & Macken, 1993; Morrison et al., 2016), the distracter familiarity effect should not emerge. However, this task is sensitive to attentional diversion effects (Hughes et al., 2007). For Study IV, all participants were required to recall the missing-item (from visually-presented series of 8 digits 1-9 during familiar and unfamiliar irrelevant melodies [without lyrics] identical to Study III, and quiet). Results were clear-cut. Irrelevant familiar melody produced poorer missing-item recall performance, with less missing-items correctly identified, than during irrelevant unfamiliar melody or quiet. As such, findings from this non-seriation task support an attentional diversion account (Marsh et al., 2018) over an interference-by-process account (Jones & Tremblay, 2000) of auditory distraction by familiar melody.

6.1 Introduction

Much theorizing about mechanisms underpinning auditory distraction has focused on serial short-term memory and its particular vulnerability to disruption via task-irrelevant sound (e.g., Beaman & Jones, 1997; Colle & Welsh, 1976; Ellermeier & Zimmer, 1997; Elliott, 2002; Hanley, 1997; Hughes & Marsh, 2017; Jones & Macken, 1993; Macken, 2014; Neath, 2000; Salamé & Baddeley, 1982; Tremblay & Jones, 1998). Serial short-term memory requirement has emerged as the quintessential task for studying mechanisms supporting short-term memory in addition to acoustic characteristics of speech that affect optimal performance of those mechanisms. Emergent from this work is that it is the acoustic variability of sound over time that renders it specifically disruptive to serial-ordering (Jones et al., 1992). A leading account of why this "changing-state" quality determines disruption of serial recall is that acoustic changes yield order cues that compete and conflict with the process of sequence-planning of vocal-motor acts in order to maintain visual stimuli from the focal task (interference-by-process; e.g., Hughes & Jones, 2001; Jones & Macken, 1993; Jones & Tremblay, 2000). While this account proposes that speech may produce more disruption of serial-recall than non-speech stimuli (Beaman & Jones, 1997; LeCompte et al., 1997; Tremblay et al., 2000; see Study III Chapter V) it attributes this to the greater acoustic complexity of speech as compared to non-speech stimuli. Furthermore, the interference-by-process view asserts that post-categorical properties of sound (in this case speech) should play no role in the disruption it produces to serial short-term memory (Jones, 1999). Post-categorical properties of speech include its semanticity (meaning) and phonology, and the impotency of these properties in determining disruption have paved the foundation of the perceptual-gestural view wherein emphasis is on the action of peripheral perceptual motor process in accounting for serial short-term memory performance (e.g., Hughes et al., 2016; Hughes & Marsh, 2017; Hughes et al., 2009; Jones et al., 2004, 2006).

Interest in the current study flows from findings of Study III (digit-serial recall, Chapter V). Consistent with an interference-by-process view was that unfamiliar and familiar speech, and unfamiliar and familiar lyrics combined with melody, were not differentially disruptive of short-term serial recall (Hughes & Jones, 2005, Jones & Tremblay, 2000). Crucially, the surprising result from Study III was that a distracter familiarity effect occurred between familiar and unfamiliar instrumental melody. (This replicated the familiarity effect shown in Study II, lyrics recall.) Crucially, however, no main effect of familiarity was evidenced when irrelevant sound contained spoken or sung-lyrics. Therefore, as melody familiarity became unimportant when combined with lyrics, this suggests lyrics overruled any disruptive impact of familiar melody. Interestingly, a distracter familiarity effect was not observed in Study I (humming) where seriation process conflict was, arguably, not as obvious. As a consequence of these findings, Study IV will focus on familiar and unfamiliar melody distraction in order to explore whether the familiarity effect can be replicated with a task thought not to require serial rehearsal.

One possible reason as to why familiar melody without associated lyrics produced more impairment of serial recall tasks than unfamiliar melody without lyrics is that pre-attentive processing of familiar melody yields stronger serial order cues that are aided by schema-driven processing (Bregman, 1990). An alternative view is that familiar against unfamiliar melody elicits attentional diversion much like emotional valence (Marsh et al., 2018) and arguably, own-name effects (Röer et al., 2013). This current study aims to cast light on the true nature of the familiar melody effect by assessing impact of familiar *versus* unfamiliar distracter melodies on a task devoid of serial recall, the missing-item task, that is sensitive to disruption via properties of sound that induce attentional diversion (e.g., emotional valence of speech distracters [Marsh et al., 2018] and unexpected violations of acoustic regularities [Hughes et al., 2007]). Establishing that the missing-item task is invulnerable to the effect of distracter familiarity would support the perceptual-gestural account while finding a disruptive effect of familiarity would support an attentional diversion account.

To reiterate here, serial recall is disrupted by the presence, within the sound stream, of perceptually segmentable elements that differ from their immediate predecessor. Thus, a sequence of repeated items (“L, L, L, L, L”) produces much less disruption (if any) than a sequence of acoustically changing items (“R, D, K, W, L”; Jones et al., 1992; Jones & Macken, 1993). That pre-categorical, acoustic, properties of speech are responsible for disrupting serial recall is buttressed by findings that it is susceptible to disruption by non-speech sounds such as tones, (Divin et al., 2001; Elliot, 2002; Jones & Macken, 1993; Sörqvist, 2010) and instrumental music (Klatte et al., 1995; Perham & Vizard, 2011; Salamé & Baddeley, 1989; Schlittmeier et al., 2008; also see Study III). A lynchpin of support for the interference-by-process view is that irrelevant sound does not disrupt serial short-term memory, a component of the perceptual-gestural account of short-term memory (Hughes et al., 2016; Hughes & Marsh, 2017; Jones et al., 2004): A gestural account eschews the notion that serial short-term memory performance is supported by a dedicated short-term memory store or working memory space (Baddeley, 2007; Cowan, 1998, 2001) and instead ascribes it to action of general-purpose perceptual and motor processes.

As outlined in Chapters II and V, the perceptual-gestural view proposes that the skill of vocal-motor sequence-planning is recruited opportunistically to bind to-be-remembered items together. Under-specification of action parameters, however, means that the skill must be populated with specific content to execute an appropriate motoric response (Hommel, 2010; Neumann, 1996). To this end, presence of other extraneous sequential information can impinge on the process of assimilating required content—visual-verbal items—during cyclical execution of the motor-plan embodying to-be-remembered content. On the perceptual-gestural account it is argued that, as a by-product of obligatory processing of changing-state sound (cf. Bregman, 1990), an extraneous sequence is generated that can interfere with motor-planning processes. Unlike processing a succession of changing-sounds that generate order cues, processing of single repeated, steady-state sounds confers little, if any, disruption of serial recall (for further details see Hughes & Marsh, 2017; Jones et al., 1996; and Chapter II).

Another principle piece of evidence favouring the interference-by-process account is the specific sensitivity of particular tasks to the changing-state effect. This view holds that only tasks whose effective performance relies upon, or tends to call upon, a process of serial ordering should be influenced by the changing-state effect. For example, free recall tends to be immune to the changing-state effect (Salamé & Baddeley, 1990) unless participants use, or are free to use, a strategy of serial rehearsal (e.g., when articulatory suppression is not required; Beaman & Jones, 1998). Importantly for the current study, identifying a missing-item (e.g., the item 7) from a well-known sequence set such as digits 1-8 presented in random order (e.g., 3184652), is also invulnerable to the changing-state effect (Elliott et al., 2016; Hughes et al., 2007; Jones & Macken, 1993; see also Beaman & Jones, 1997). The missing-item task is one in which a serial order strategy tends not to be adopted (Morrison et al., 2016). However, participants' performance on the missing-item task is susceptible to the changing-state effect if they self-report using a rehearsal strategy (Hughes & Marsh, 2019), or if visual memoranda comprised items that a participant had no long-term sequential order for (e.g., 'buildings for religious services' as compared with 'days of the week'), or if they were not learned in a fixed order (e.g., alphabetical; see Beaman & Jones, 1997).

In light of the foregoing, the findings in Study III, that familiar distracter melodies were more potent to disrupt than unfamiliar distracter melodies, would appear altogether inconsistent with the interference-by-process account (Hughes et al., 2009; Jones, 1993; Jones & Tremblay, 2000). This is because such an effect must go beyond simplistic conceptualization of the changing-state effect for which token-to-token acoustic changes yield order cues and drive the disruption effect. Moreover, it implies some post-categorical processing (e.g., identification of a melody). However, for this current study, the missing-item task does not require post-categorical processing, semantic or phonological, rendering any such interference by this process defunct.

For detection of a familiar melody to occur, the first two or three elements of a temporarily unfolding pattern must be detected. Processing of these initial elements of a familiar melody therefore, likely activates a schema and therefore entails schema-driven, as compared with primitive, grouping, whereby concepts acquired through experience are applied to acoustic stimuli as compared to grouping based on incoming acoustic data (Bregman, 1990) that drives the changing-state effect (Jones, 1993; Jones & Tremblay, 2000). A crucial question concerns how such schema-driven processing, as invoked by a familiar melody, gives rise to greater disruption of serial-recall than an unfamiliar melody, and whether this can be reconciled with the interference-by-process account. One possibility is that exposure to a familiar melody gives rise to a predictive model (Bell, Röer, Marsh, Storch, & Buchner, 2017; Hughes & Marsh, 2017, 2019; Röer et al., 2015) concerning expectancies of the forthcoming serial order of melodic elements. Furthermore, although speculative, it is possible that these internally-generated serial order cues are stronger than those computed from acoustic changes via primitive streaming, much like the generation effect for item interference (Burns, 1990; Mulligan, 2002; Nairne, Reigler, & Serra, 1991). If this were the case, then greater interference-by-process would be expected from familiar as compared to unfamiliar melody, since the strength of order cues within a to-be-ignored stream governs object formation and serial order (Hughes, 2014; Hughes et al., 2009).

A second possibility is that any attentional disruption produced by familiar, compared to unfamiliar, distracter melodies is attributable to attentional diversion. That is to say attention may be momentarily disengaged from the focal task due to the presence of familiar melodic auditory stimuli (Hughes et al., 2007; Marsh et al., 2018; Vachon et al., 2017). Attentional diversion effects can be fractionated into those that arise due to *specific attentional capture,* such as when the sound's particular content empowers its capability to divert attention (e.g., sound of one's own name amidst general conversation), and *aspecific attentional capture,* whereby a sound captures attention because of the context in which it occurs (e.g., an unexpected alarm bell in a quiet environment; for further discussion see Eimer et al., 1996). Thus, attentional diversion attributable to a stimulus can arise due to the context of the stimulus itself or purely the environmental context within which the stimulus is embedded. Since familiar distracter melodies in Study III did not contain elements that violated context (such as an out of key deviant note) and yet produced more disruption than unfamiliar distracter melodies, the effect, if attributable to attentional diversion, would appear to be due to specific attentional capture. Crucially, unlike the changing-state effect that is restricted to tasks requiring serial order processing (e.g., Hughes et al., 2007; Jones & Macken, 1993), both specific and aspecific attentional diversion effects are observed for tasks that do not necessitate or require serial order processing (e.g., Hughes et al., 2007; Parmentier et al., 2008; Vachon et al., 2017). For example, performance of the missing-item task, and the serial recall task are both impaired by an auditory item that violates a preceding pattern of stimulation (aspecific attention capture), and the emotional valence of auditory distracter words (post-categorical context; Marsh et al., 2018).

At first glance it might not appear obvious how a familiar melody (as observed in Study III) could disrupt seriation more than an unfamiliar melody in the context of the interference-by process account (Jones & Tremblay, 2000). Since they were rhythmically and harmonically matched, with step/leap pitch contour equivalence, familiar and unfamiliar melodies should have possessed comparable changing-state sound properties. However, based on previous literature, it is possible to envisage a scenario in which familiar irrelevant sound sequences could generate greater interference-by-process than unfamiliar irrelevant sound sequences. One possibility is that there exists an arousal-based attentional capture. In support of this contention, pupil dilation has shown to be greater for familiar melodies than for unfamiliar melodies, suggesting an enhanced arousal response follows exposure to familiar melodies (Weiss, Trehub, Schellenberg, & Habashi, 2016). Weiss et al. (2016) also identified a greater dilation for familiar melodies that were presented by voice (sung la, la) as compared to by piano. However, it should be noted that familiarity for this task was episodically-determined with a familiar stimulus classified as such by just two exposures to a previous unfamiliar folk melody. While arousal may serve as an explanation for differences observed, evocation of emotional engagement is reliant on long-term memory associations.

Limited neuropsychological studies have focused on music familiarity, but ERP studies have identified that familiar music activates the brain’s right hemisphere whereas unfamiliar melody activates the left hemisphere (Plailly et al., 2007).

Furthermore, it has been demonstrated, through an electrodermal activity (EDA) study, that familiarity can play a part in emotional arousal (Van Den Bosch, Salimpoor, & Zatorre, 2013). The idea that arousal-based attention capture arises because familiar melody connects with neural areas responsible for processing the affect is, therefore, a consideration. There is also evidence that familiar melody activates brain areas that are responsible for sequence planning (Halpern & Zattore, 1999; Koelsch et al., 2009). Supplementary motor areas (SMA), crucial for various facets of motor behaviour, have been shown to be activated in music listening and in real and imaginary performance (Lima et al., 2016; Tanaka & Kirono, 2017). The SMA is also heavily involved in articulation and motor-planning for serial short-term memory (e.g., Halpern & Zatorre, 1999; Hughes et al., 2009; Jones et al., 2006; Jones et al., 2004; Koelsch et al., 2009). For example, according to the perceptual-gestural account (Hughes et al., 2009; Jones et al., 2004, 2006), verbal items are assembled in articulatory form, the speech-planning mechanism then co-opted to compensate for loss of linguistic aspects, such as syntax and semantics, removed from serial recall (see Chapter I). Therefore, motor-output organisation that arguably draws upon the SMA is crucial for serial short-term memory (e.g., Hughes & Jones, 2005; Woodward et al., 2008). If familiar melody activates the SMA to a greater extent than unfamiliar melody, then interference with sub-vocal motor-output organisation required for serial recall would be greater.

If familiar melody seizes control of, or recruits, motor-planning systems more than unfamiliar melody then it would be expected that familiar melody would impair serial recall (Study III), but not the missing-item task (Study IV) which does not require such sequential motor-planning (serial rehearsal). If this was the case, then any additional disruption produced by familiar, as compared to unfamiliar, melody would constitute a more specific interference-by-process effect. Alternatively, if any additional disruption that familiar melodies produce over unfamiliar melodies is due to attentional diversion, then any familiar melody distracter effect should be found in the context of the missing-item task (Study IV). Since attentional diversion effects, such as the deviation effect (Hughes et al., 2007; Vachon et al., 2017), the distracter valence effect (Marsh et al., 2018), and sudden prosody deviations in irrelevant speech (Kattner & Ellermeier, 2018) disrupt performance on both serial recall and the missing-item task, then it should be considered that attentional diversion may also be produced by a familiar as compared to an unfamiliar melody distracter.

6.1.1 The missing-item task.

For Study IV evidence was sought to determine the underlying nature of the familiar melody distracter effect observed in Study III. The attentional diversion account holds that the familiarity effect should obtain regardless of whether serial order processing is adopted to meet the demands of the short-term memory task. On the contrary, if the distracter familiarity effect is attributable to stronger order cues, then the interference-by-process account holds that only tasks where serial order processing is necessary or recruited should be vulnerable to the familiarity effect. Therefore, a missing-item task, for which serial ordering strategies are quite infrequently self-reported, was adopted for Study IV (Hughes & Marsh, 2019; Morrison et al., 2016).

The presentation condition of the missing-item task is similar to the serial recall task. Participants are sequentially presented with a set of digits (e.g., 8) from a set (e.g., 1-9). Instead of recalling all items in serial order following the offset of the last item, participants are required to identify the item that was not presented from the set, the missing-item. For example, number 2 missing from the list 47519836 (Buschke, 1963; Klapp et al., 1983). Unlike serial recall, retention of item order is not necessary for missing-item identification. Moreover, two central findings support the notion that participants do not generally adopt a serial ordering strategy for this task. Unlike order-recall tasks, such as serial recall, the missing-item task is invulnerable to disruptive effects of concurrent articulation on performance (Beaman & Jones, 1997; Klapp et al., 1983). Such articulatory suppression—which occupies the mechanism driving inner/covert speech-planning—is typically assumed to prevent or at least impair serial rehearsal processes (Baddeley, 2007; Jones et al., 2006; Murray, 1968). It should be noted here that an effect also arguably located in the serial rehearsal process—the talker variability effect—is not obtained in the missing-item task (Hughes et al., 2011). Furthermore, in a recent analysis of self-reported strategy-use across a number of different short-term/working memory tasks, serial rehearsal was only reported by 25% of participants performing the missing-item task (Morrison et al., 2016). The most common strategy reported for this task is one of "checking off". Here, participants report checking off each item as it appears against a representation of fixed ordinal sequence. Recognition of which item was not checked off then forms the basis of missing-item identification (Beaman & Jones, 1997; Buschke & Hinrichs, 1968; Humphreys & Schwartz, 1971; Morrison et al., 2016). There is good evidence that the missing-item task can be performed without recourse to serial order processing. Prior studies have shown the changing-state effect, and therefore ordering of acoustic change within a task-irrelevant sound, has little, if any, effect on the missing-item task (Beaman

& Jones, 1997; Elliott et al., 2016; Hughes et al., 2007; Jones & Macken, 1993), unless a serial order strategy is reported (Hughes & Marsh, 2019). Contrary to this, the deviation effect (aspecific attentional diversion; Hughes et al., 2007; Vachon et al., 2017), and emotional valence (specific attentional diversion; Marsh et al., 2018) effects of auditory distraction are readily observed in the missing-item task.

The rationale of comparing performance in the presence of auditory distracters in the context of a serial-digit recall task (Study III), with a missing-item task (Study IV), was to ascertain if the distracter familiarity effect discovered in Study III (whereby familiar melody disrupts serial recall performance to a greater extent than unfamiliar melody) is due to familiar sounds causing attentional diversion or interference-by-process. Comparing the impact of a particular distracter sequence on a serial recall task with a missing-item task is emerging as a key theoretical device to examine mechanism(s) underpinning distraction of focal task by sound (Dorsi, Viswanathan, Rosenblum, & Dias, 2018; Kattner & Ellermeier, 2018; Marsh et al., 2018). If the distracter sequence impairs performance of both serial recall and missing-item tasks, then it can be classified as an attentional capture effect and therefore qualitatively distinct from the classic changing-state effect (interference-by-process). Thus, it was predicted that the missing-item task should be invulnerable to the distracter familiarity effect if it was produced by acoustic changes within the sound/stronger order encoding (interference-by-process; Jones & Tremblay, 2000). However, finding a distracter familiarity effect for the missing-item task in addition to the serial recall task (Study III Chapter V) would bear evidence for a specific attentional diversion effect (Hughes et al., 2007; Vachon et al., 2017).

6.2 Experiment 4: Method

6.2.1 Participants.

One hundred and six participants were recruited for this Study from UCLan and community groups in Lancashire. While an initial scrutiny revealed no floor effect, five participants scored the maximum (1) in the quiet and familiar melody distracter conditions. Five participants were therefore excluded due to performing at ceiling and were replaced. Henceforth, all data embodies five replacement participants. Participants comprised 29 Men and 77 Women aged 18-85 (M = 43.62, SD = 23.04[52]). All participants had an opportunity to read the information sheet and to ask the

[52] The mean was comparable with Study III (M = 46.55, SD =22.68).

researcher any questions before being asked to complete consent and demographic forms.

Cognition.

In keeping with previous experiments in this series, 44 participants aged over 50 undertook Addenbrooke's Cognitive Examination Revised (ACE-R) which also gives mini-mental state examination (MMSE) information. Scores ranged between 84-99 (M = 94.23, SD = 3.26) with an MMSE range of between 24-30 (M = 28.86, SD = 1.19) which indicates positive cognitive functioning[53]. No participants scored <82. A visual acuity test for participants aged 50+ confirmed all had acceptable vision (task font 16), and a hearing loss test, conducted via an audiometer, recorded normal hearing levels[54].

Musical culture.

For the current study 100% of participants indicated a 'Western' musical culture on their individual response sheet, with dual cultures identified by 4.2%.

Musical training and participation.

Participants also indicated on their individual demographic response sheet their years of musical training (0-8+ years), and current participation in a choir or instrumental group. An analysis of variance (ANOVA), comparing recall scores with years of musical training, indicated that all were non-significant on recall accuracy and there were no significant interactions[55]. There was no between-participant effect.

6.2.2 Design.

Regarding the auditory conditions, contrary to Study I-III, only familiar and unfamiliar melody distracter material was delivered for this current study as a within-participants design. This decision resulted from the familiarity effect identified in Study III observed only for the Melody Sound Type Group. The single Sound Condition distraction variable was classified into three levels, quiet (no distracter), unfamiliar melody, and familiar melody. A shortened duration for each sound file to 10-seconds, used in Study III was adopted for Study IV. Two blocks of three conditions yielded six orders of presentation, as shown in Table 6.1. In order to address repeated measure

[53] A cut off <88 gives 0.94 sensitivity for dementia, <82 gives 0.84. MMSE >24 no cognitive impairment.
[54] Normal hearing established at 0-25 dB, some mild hearing loss at whisper level recorded at 20-40 dB.
[55] Sound: $F(2, 186) = 1.183$, $MSE = .015$, $p = .309$, $\eta_p^2 = .013$
Sound × Training: $F(6, 186) = 1.442$, $MSE = .019$, $p = .201$, $\eta_p^2 = .044$
Sound × Choir: $F(2, 186) = 1.421$, $MSE = .018$, $p = .244$, $\eta_p^2 = .015$
Sound × Instrument: $F(2, 186) = 2.146$, $MSE = .028$, $p = .120$, $\eta_p^2 = .023$

unsystematic variation, orders were counterbalanced, each order delivered to eighteen participants.

Table 6.1

Order of condition presentation for participants

	Sound Condition	Order	Trials						Order	Trials					
1	quiet	**1**	1	3	2	1	3	2	**4**	3	2	1	3	2	1
2	unfamiliar melody	**2**	1	2	3	1	2	3	**5**	2	3	1	2	3	1
3	familiar melody	**3**	3	1	2	3	1	2	**6**	2	1	3	2	1	3

6.2.3 Materials and apparatus.

Ten familiar distracter melodies, taken from nursery rhymes or traditional songs that all had associated lyrics, together with ten unfamiliar matched melody distracters, were the same as those delivered in the Melody Sound Condition for Study I-III with one modification: An additional set of ten familiar melodies (without associated lyrics), and an additional ten unfamiliar matched melodies (without lyrics) were included for this task, creating a second block of trials of identical structure. The rationale for including additional trials here was to yield a richer set of data since an identified missing-item, right or wrong, creates only one data point per trial, as compared to eight data points per trial produced from the serial-digit recall task. To enable a better comparison with results from previous studies, additional melodies were always presented in the second block (such that, if block 2 melodies were ineffective, previously used melodies could be contrasted with foregoing studies without the confound of block order). Sound files were all compiled using Avid Sibelius7 music notation software, all sound stimuli having the same format and properties as for previous studies in this series. The task was developed using the E-Prime 2 (sp1) Psychology software tool and conducted via a laptop computer with the sound level averaging 62dB (A).

6.2.4 Procedure.

Following five practice trials in quiet conditions participants were visually presented with series of eight single digits drawn from digits 1-9. Instructions directed participants to identify the missing digit. Each digit was presented once in each of ten trials for each of six condition blocks. Participants were informed they should ignore any sounds being played through headphones (Sennheiser HD201) at any time during the task. Onset of a to-be-ignored sound in the unfamiliar and familiar melody conditions was simultaneous with each visual digit pattern presentation. To ensure standardisation of instructions to participants instructions to “Begin Trial” were consistent for each movement through the program. Prior to beginning each trial a visual orienting cue + was shown in the centre of the screen to focus participants' attention on that area of the screen where item information was to be shown. Digits were presented consecutively at a rate of 1 per second (800-millisecond on, 200-millisecond off) in 72-point equidistant black Monaco font on a white background. From a viewing distance of 45 centimeters each number subtended a vertical visual angle of 1.49° and a horizontal angle of 0.92°. Following a 200-millisecond retention interval at the end of each trial, all nine of the digits 1-9 were presented on the computer screen with instruction to double click the mouse pointer on the digit that was missing from the original trial sequence. There was no time-limit on recall. A 30-second rest between experimental conditions was permitted, if needed, this proposal indicated via the computer screen.

At the end of the experimental trials, participants were instructed, via computer screen, to indicate a strategy used to complete the missing item task. These were based on strategies suggested by Morrison et al. (2016). Statements such as, for example, "I silently repeated the items" covered approaches related to rehearsal, grouping, semantics, association, look, imagery, sound, concentration, familiarity, and checklist. In addition, participants could indicate if an unidentified strategy had been used or if they did not understand the instruction/demands (Appendix E).

To control for prior knowledge of familiar and unfamiliar distracter melodies (to establish a baseline), once all trials had been completed, participants were played the 40 melodies again. Participants were requested to press the "Y" key if the melody was familiar and the "N" key if not familiar. They were then instructed to indicate their confidence in that decision by clicking on a point across a continuum ranging from "Not confident" to "Totally confident". This procedure mirrored that used for Study III. The experiment took approximately 45 minutes to complete.

6.2.5 Assessment.

For Study IV, the main purpose was to identify the effect of an unfamiliar or familiar melody distracter on performance accuracy in a non-serial, missing-item task. This would be achieved by comparing means for each of three levels of Sound Condition variable: quiet, unfamiliar melody, and familiar melody. Accuracy for identification of the missing digit was recorded by the E-Prime program and from the generated data files, mean accuracy across same-trial types were computed from block 1 and block 2 for analysis. Similarly, mean scores for Recognition data and confidence in recognition decisions were also extracted from data files for analysis.

6.3 Results

Overall scores in quiet conditions for the missing-item task were shown to be the highest (maximum 1). Performance decreased in the unfamiliar melody condition and there was a further decrement to scores for the familiar melody condition. Interestingly, although patterns were consistent across blocks 1 and 2, there was an increase in accuracy across all conditions in block 2, possibly suggestive of a more highly developed consistent strategy honed through practice, see Table 6.2. However, for both blocks, scores in quiet condition were highest with lowest scores obtained in the familiar melody condition, see Figure 11.

Table 6.2

Missing-item mean scores overall, and within blocks, across Sound Conditions

Sound Condition	Overall Mean	*SE*	Block 1	*SE*	Block 2	*SE*
quiet	.693 *(.187)*	.018	.675 *(.220)*	.021	.708 *(.211)*	.021
unfamiliar	.675 *(.199)*	.019	.665 *(.208)*	.020	.685 *(.239)*	.023
familiar	.631 *(.202)*	.020	.618 *(.235)*	.023	.643 *(.233)*	.023

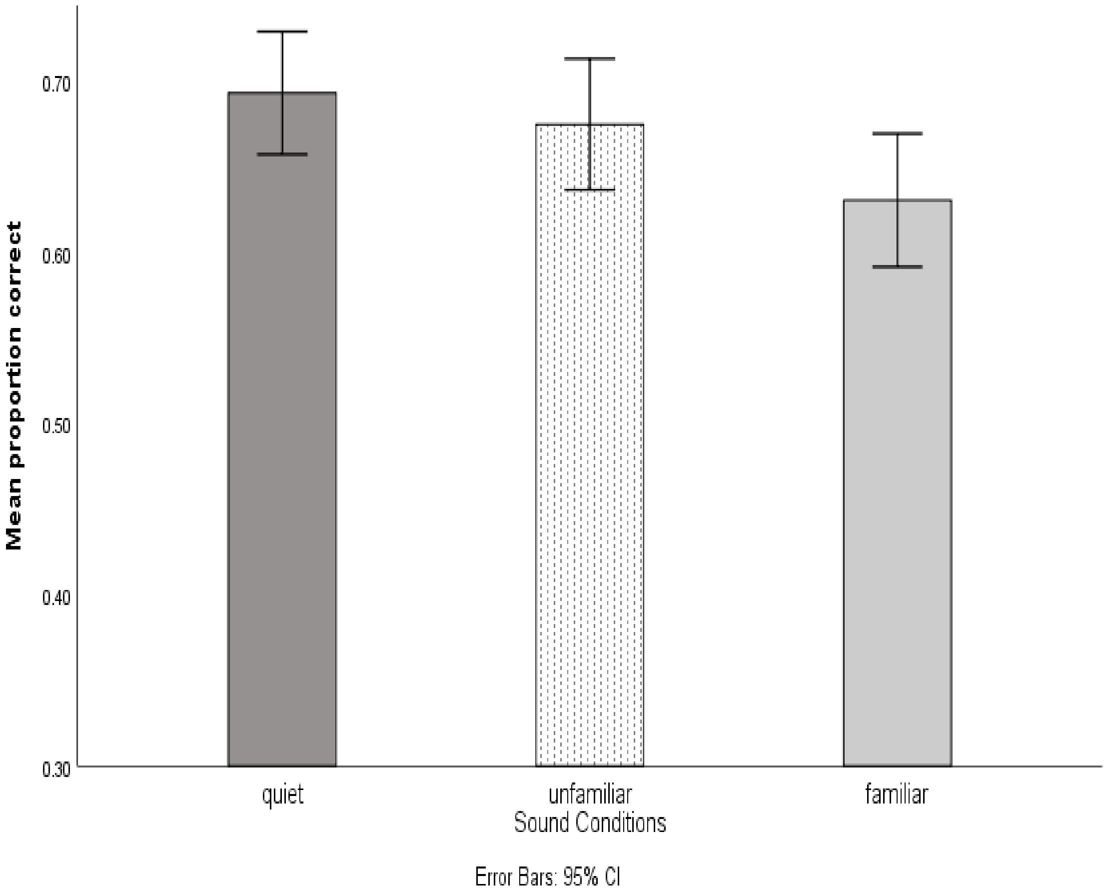


Figure 11. Missing-item overall mean recall scores across Sound Conditions.

A repeated measures ANOVA showed that overall there was a significant main effect for Sound Condition, $F(2, 210) = 8.544$, $MSE = .013$, $p < .001$, $\eta_p^2 = .075$. Pairwise comparisons demonstrated better performance for quiet compared with familiar melody ($MD = .063$, $SE = .016$, $p < .001$; 95% CI [.030, .095]), and performance in the unfamiliar melody condition was superior to performance in the familiar melody condition ($MD = .044$, $SE = .016$, $p = .007$; 95% CI [-.013, .076]), thereby demonstrating a familiarity effect. There was no significant difference between quiet and unfamiliar melody ($MD = .018$, $SE = .014$, $p = .200$; 95% CI [-.010, .047]) suggesting that the presence of sound *per se* did not produce significant disruption on the missing-item task.

A further repeated measures ANOVA was conducted to compare accuracy across the two blocks of trials as a function of Sound Condition. This revealed a main effect of Sound Condition, $F(2, 210) = 8.375$, $MSE = .025$, $p < .001$, $\eta_p^2 = .074$, and a main effect of Block, $F(1, 105) = 3.991$, $MSE = .027$, $p = .048$, $\eta_p^2 = .037$, that owed to performance being slightly better for the second as compared to the first block. There was no interaction between Sound Condition and Block, $F(2, 210) = .101$, $MSE = .023$, $p = .904$, $\eta_p^2 = .001$.

Participants self-reported a range of strategies for determining the missing digit. The most frequently used related to grouping, rehearsal, and checklist, as shown in Table 6.3[56]. A preponderance for a grouping strategy, although not a prerequisite, possibly indicated that some serial rehearsal was indeed employed to determine the missing-item.

Table 6.3

Participant self-reported strategy used to determine the missing item

Strategy	Approaches	% used	Cumulative %
c	checklist	26.6	26.6
d	rehearsal	25.7	52.3
e	grouping	27.5	79.8
f g i j	semantics/visual/meaning/concentration	18.4	98.2
a b h k	familiarity/look/sound/imagery/semantics	1.8	100

[56] Strategies identified in Appendix E.

In order to ascertain whether the familiarity effect would appear for participants using a rehearsal strategy, groups were then divided into 58 rehearsers (those using serial rehearsal [d] or grouping [e]) and 48 non-rehearsers (all other strategies combined). The means for these two groups self-reported differential use on rehearsal were then compared. Results of the ensuing Sound Condition × Strategy (non-seriation *vs*. seriation) ANOVA revealed no between-participant main effect of Self-Reported Strategy, $F(1, 104) = .000$, $MSE = .090$, $p = .999$, $\eta_p^2 = .000$, and no interaction between Sound Condition and Strategy, $F(2, 208) = .697$, $MSE = .013$, $p = .499$, $\eta_p^2 = .007$, suggesting that the impact of distracters on recall accuracy was not influenced by the Self-Reported Strategy used to remember the missing digit.

6.3.1 Recognition and confidence test.

Total scores from the Recognition test indicated a mean of 14.37/20 (*SD* = 3.031; *SE* = .294) correct identifications of familiar melodies (Hits), and 16.92/20 (*SD* = 3.464; *SE* = .336) correct rejections (CR) of unfamiliar matched melodies. This yielded a hit-rate of .718 (*SD* = .152; *SE* = .015), and a false alarm rate (FA) of .154 (*SD* = .173; *SE* = .017), see Figure 12. D'prime and criterion shift (c) were also computed for recognition scores. Data for 28 participants were excluded from this analysis since those participants did not commit a miss response (1), a false alarm response (26) or both (1) barring computation of both measures. The d'prime measure (of recognition sensitivity/discriminability) was 1.544 (*SD* = .768; *SE* = .087). The criterion shift score was .192 (*SD* = .382; *SE* = .043). This suggests that participants[57] were able to discriminate unfamiliar from familiar melodies and did so conservatively, requiring considerable evidence before deciding a melody was familiar. This is frequently observed when distracters are similar, compared with dissimilar, to targets (e.g., Benjamin & Bawa, 2004).

[57] $n = 78$

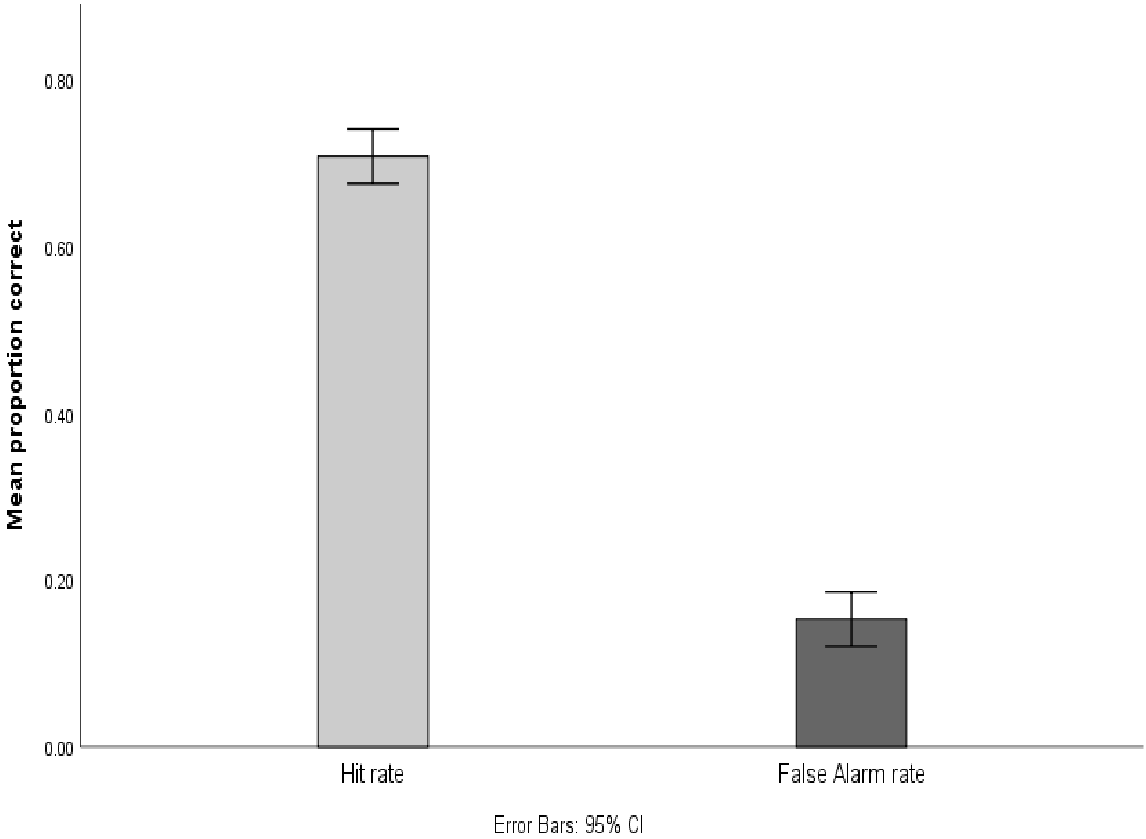


Figure 12. Comparison of Hit rate and False Alarm rate for the Recognition test.

For each of the 40 melodies, participants also rated confidence in their response across a continuum from 0-750 as shown in Table 6.4. Confidence was lowest for the incorrect unfamiliar matched melodies that were endorsed (false alarms).

Table 6.4

Confidence scores for missing-item familiar and unfamiliar melody recognition

Confidence	Participants	Mean *(SD)*	SE
familiar melody - Hits	78	674.88 *(59.52)*	6.74
familiar melody - Miss	78	526.55 *(172.07)*	19.48
unfamiliar melody - CR	78	538.11 *(161.99)*	18.34
unfamiliar melody - FA	78	479.40 *(165.21)*	18.71

As for d'prime and c computations, participants who did not make a miss or false alarm were excluded for the ensuing one-way repeated measures ANOVA for confidence level as a function of Decision Type (Hit, Miss, Correct Rejection, False Alarm). A main effect of Decision Type was revealed[58], $F(2.083, 160.369) = 39.487$,

[58] Huynh-Feldt reported following sphericity violation $\varepsilon = .694$

MSE = 20088.596, *p* < .001, η_p^2 = .339. Follow-up pairwise comparisons revealed that participants were more confident in hit decisions than miss (*MD* =148.328, *SE* = 18.363, *p* < .001; 95% CI [111.763, 184.893]), correct rejection (*MD* = 136.771, *SE* = 17.147, *p* < .001; 95% CI [102.628, 170.914]), and false alarm (*MD* = 195.478, *SE* = 18.341, *p* < .001; 95% CI [158.957, 231.999]) decisions. Furthermore, confidence in correct rejection decisions was higher than confidence for false alarms (*MD* = 58.708, *SE* = 22.280, *p* = .010; 95% CI [14.342, 103.074]), and confidence in misses was higher than confidence in false alarms (*MD* = 47.150, *SE* = -23.881, *p* = .052; 95% CI [-.403, -94.703; although this difference approached significance]). However, confidence levels in miss and correct rejection decisions did not differ (*p* = .277). Therefore, participants were most confident when endorsing a familiar melody as familiar (hit) and least confidence when they responded that an unfamiliar melody was familiar (false alarm). Confidence levels for responding that a familiar melody was unfamiliar (miss) and an unfamiliar melody was unfamiliar (correct rejection) were comparable.

6.3.2 Response time.

For the current missing-item task, response time was not measured by the E-Prime program. The missing-item task only requires a single response and thus response times are less informative than those for melody/lyrics production (Study I-II), or serial recall (Study III) that require assembly and planning of sequential output (Hughes et al., 2009).

6.4 Discussion

The aim of Study IV was to address mechanisms underpinning the distracter familiarity effect—greater disruption to short term-memory performance attributable to a familiar against unfamiliar melody—observed in Study III. An interference-by-process view (e.g., Jones & Tremblay, 2000; Marsh et al., 2009) was pitched against an attentional diversion (e.g., Marsh et al., 2018) account of the distracter familiarity effect using a missing-item task that can be accomplished without recourse to maintenance of items in serial order via vocal-motor sequence planning (e.g., Hughes & Marsh, 2019; Jones & Macken, 1993). It was reasoned that if the distracter familiarity effect was attributable to a clash between serially-ordering elements of irrelevant sound and use of serial rehearsal for the focal task, then it should fail to emerge on the missing item task. Similarly, it was deduced that if the distracter familiarity effect were attributable to specific attention capture, it would indeed emerge on the missing-item task, since attentional capture effects are not task-process dependent (Hughes et al., 2007; Marsh et al., 2018). This discussion provides a synthesis of results from the non-serial missing-item task in relation to possible theoretical mechanisms at play in producing a distracter familiarity effect.

6.4.1 Overall results.

Results for Study IV were unequivocal: the missing-item task was vulnerable to the distracter familiarity effect. Furthermore, a distracter familiarity effect was equally observed for participants self-reporting a non-seriation strategy and a seriation strategy as reflected in the absence of a Sound-Condition × Self-Reported Strategy interaction. Further analysis—by means of a recognition test—revealed that participants adjudged that familiar as compared to unfamiliar melodies were indeed familiar, supporting validity of the stimuli used.

6.4.2 Interference-by-process *versus* attentional capture.

That the missing-item task was vulnerable to the distracter familiarity effect—regardless of whether the self-reported strategy was seriation-or non-seriation-based—supports the idea that the mechanism undergirding the effect is attentional capture (Marsh et al., 2018) rather than interference-by-process (Jones & Tremblay, 2000). Use of a missing-item task with unfamiliar and familiar distracter melody accompaniment yields results that offer theoretical precision which use of the serial recall task alone could not deliver. It was initially surprising that familiar melodies disrupted short-term memory to a greater extent than unfamiliar melodies for serial order in Study III. However, this distracter familiarity effect could have been reconciled with the interference-by-process account (Jones & Tremblay, 2000) by which serial rehearsal of to-be-remembered items is vulnerable to pre-attentive processing of the serial order of changing acoustic elements within the irrelevant auditory sequence (Jones, 1992). An initial perplexity was that changing-state properties of familiar and unfamiliar melodies (e.g., changing pitch movements/unfolding acoustic variability) were equated and so a factor other than mere acoustic variability would be required to account for greater disruption familiar as compared to unfamiliar melodies produced to serial recall. One possibility was that familiar melodies, being overlearned, may activate serial-order representations—such as a competing sub-vocal motor plan (e.g., Lima et al., 2016)—to a greater degree than unfamiliar melody distracters, endowing them with a superior propensity to clash with sub-vocal motor output organization underpinning the serial rehearsal process, thereby exacerbating interference-by-process on the serial recall task. This interference-by-process account of the distracter familiarity effect necessarily assumes that it is contingent on recruitment of a seriation process on the focal task (seriation: Jones, 1993; Jones & Tremblay, 2000).

However, results from Study IV provide clear-cut evidence that the distracter familiarity effect was produced by (specific) attentional capture as compared with interference-by-process. Consistent with the attentional capture view, while at odds with the interference-by-process view, Study IV demonstrated that the distracter familiarity effect occurred for a missing-item task, widely thought to involve non-seriation processes (e.g., Beaman & Jones, 1997; Buschke & Hinrichs, 1968). Moreover, further evidence in favour of the attentional capture view (cf. Marsh et al., 2018) was gleaned from the fact that additional familiar over unfamiliar melody disruption occurred regardless of whether participants self-reported a seriation or non-seriation-based strategy. Thus, the distracter familiarity effect is task-process insensitive.

While evidence from Study IV supports an attentional capture view, there are some issues in need of interpretation. For example, the lack of an interaction between Sound Condition and Self-Reported Strategy in Study IV might be considered surprising since one would expect unfamiliar distracters to be disruptive for participants engaging in a seriation strategy for which a changing-state effect (interference-by-process) should be observed (Jones & Tremblay, 2000; see also Study III). However, the overall pattern is for familiar melody distracters to impair performance more than unfamiliar melody distracters and quiet, but for unfamiliar melody distracters to produce no more disruption relative to quiet. On the face of it, self-reported strategies appear to differentiate between strategies used on tasks. For example, self-reported use of a checklist strategy is common for the missing-item task (Morrison et al., 2016). However, it is of no guarantee that all participants interpreted the strategy questionnaire in the same manner. Furthermore, strategy labels may be too broad to capture intricacies of differences within a strategy with “rehearsal” encompassing maintenance repetition (repeating one item at a time) and cumulative repetition, whereby two or more items are covertly repeated in succession, as would occur for serial rehearsal (cf. Lehmann & Hasselhorn, 2007). A further way to address the pattern of irrelevant sound effects as a function of task-process is to directly compare results of Study IV with results from the same Sound Conditions of Study III for which serial recall was required to be undertaken by participants.

6.5 Task comparison: Missing-item *versus* serial-digit

To cast further light on mechanisms underpinning the changing-state effect (as indexed by quiet and unfamiliar melody difference) an assimilation of the results of Study IV in relation to those obtained from Study III provide a clearer conceptual understanding of how different mechanisms of distraction may interact with the characteristics of the task. A comparison analysis of recall performance across the missing-item task (Study IV) and the serial recall task (Study III) provided an opportunity to probe further into the impact of unfamiliar melody on task performance across two tasks making differential demands on serial rehearsal processes. An interaction between Sound Condition and Task was predicted whereby unfamiliar melody should be more detrimental for serial recall task than a missing-item task. This is because unfamiliar melody should produce a changing-state effect whereby the pre-attentive processing of the serial order of changing acoustic elements within the melody should impair deliberate serial organisation of to-be-remembered items via the serial rehearsal process that primarily underpins serial recall, as compared with missing-item, task (Beaman & Jones, 1997; Jones & Macken, 19930.

A 2 (Sound Condition: quiet, unfamiliar) × 2 (Task: missing-item, serial recall) mixed factor ANOVA showed a significant main effect for Sound Condition, $F(1, 204) = 16.487$, $MSE = .010$, $p < .001$, $\eta_p^2 = .075$. There was no between-participants main effect of Task, $F(1, 204) = .245$, $MSE = .060$, $p = .621$, $\eta_p^2 = .001$. However, the Sound Condition × Task interaction was significant, $F(1, 204) = 4.645$, $MSE = .010$, $p = .032$, $\eta_p^2 = .022$. A follow-up simple effects analysis[59] to unpack this interaction confirmed that unfamiliar melody disrupted the serial recall task (MD =.060, SE = .014, $p < .001$; 95% CI [.033, .087]), but not the missing-item task (MD =.018, SE = .013, p = .173; 95% CI [-.008, .045]).

As expected, results shown in Figure 15, indicate unfamiliar melody only disrupts performance on the focal task that required seriation (Study III, serial-digit recall). This result is particularly supportive of the interference-by-process account, wherein it can be assumed that the disruption represents the changing-state effect (Jones & Tremblay, 2000).

[59] LSD reported as it maintains statistical power. See Chapter III, 3.2.5.

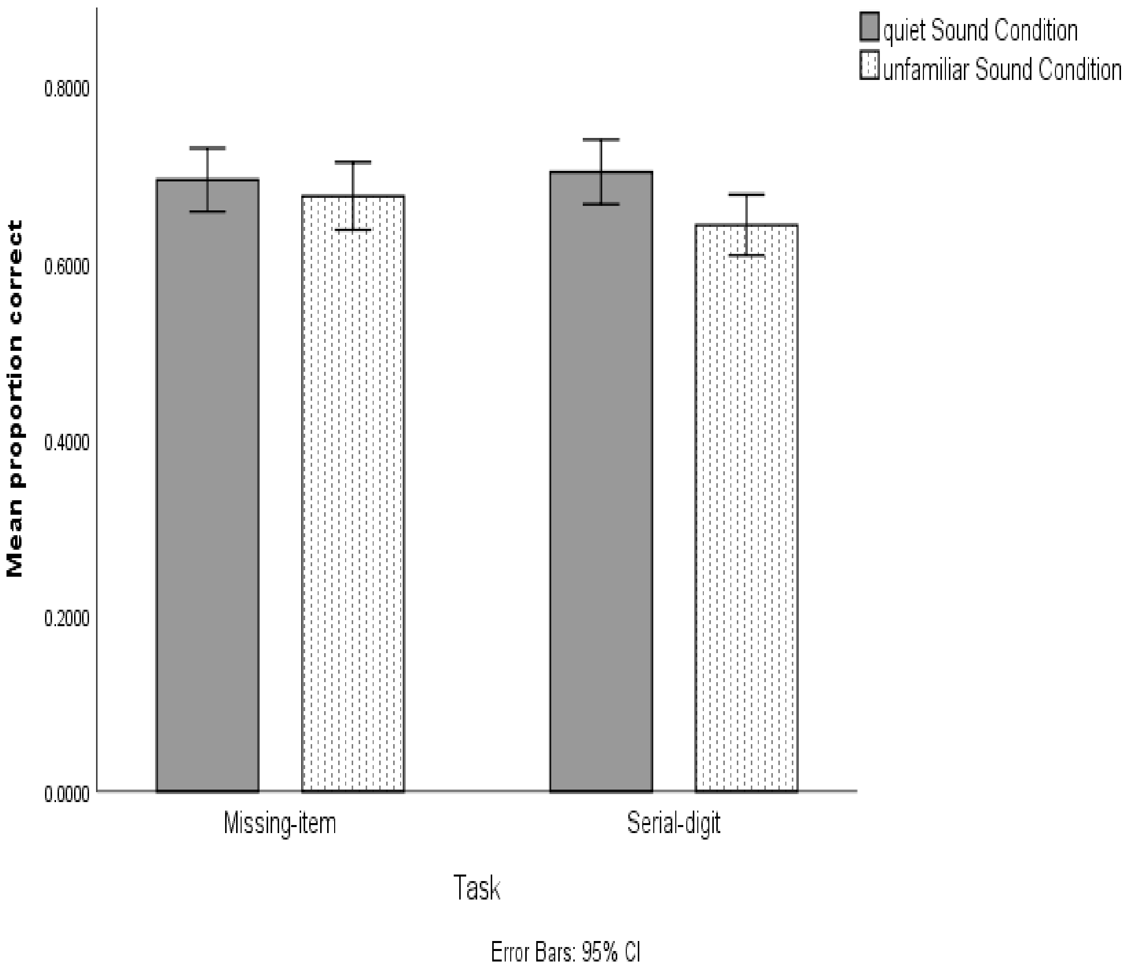


Figure 13. Mean recall scores from missing-item and serial-digit task comparison as a function of quiet and unfamiliar Sound Conditions.

6.6 A duplex account of distraction by melody?

The duplex mechanism account (Hughes et al., 2005, 2007, 2013) holds that auditory distraction in the context of short-term memory can occur via one or two qualitatively distinct mechanisms; interference-by-process and attentional capture. To recap, interference-by-process is assumed to arise because the order of changes within sound are automatically processed as part of a perceptual streaming process (Bregman, 1990) and generate order cues that interfere with deliberate sequential planning of vocal-motor acts to maintain visual items in sequence (Jones & Tremblay, 2000). This account explains why tasks that involve seriation (free recall, serial recall) are prone to greater disruption from changing-state, as compared with steady-state, sounds—the changing-state effect—while tasks that do not necessitate seriation—such as the missing-item task—are invulnerable (Hughes et al., 2007). The invulnerability of non-seriation-based tasks to the changing-state effect is a seemingly insurmountable object for unitary, attentional accounts to explain. According to the unitary account (Bell et al., 2010, 2012; Cowan, 1995; Cowan et al., 2005; Elliott, 2002), any changing state sound captures attention from a demanding focal task regardless of whether task performance requires seriation. This general failure of the unitary account to explain the process-sensitive nature of distraction emerged again in the current task-comparison analysis. Here, unfamiliar melody which was disruptive of serial recall (Study III) failed to disrupt missing-item performance (current study).

According to the duplex mechanism account attentional capture effects occur regardless of whether seriation is involved in task performance. Attentional capture can arise when the specific content of a sound yields its disruptive power (e.g., a food related sound for a hungry person; Parmentier et al., 2018), or when it violates an unexpected context (e.g., a ring-tone sounding in a quiet lecture theatre). These have been referred to as specific attentional capture effects and aspecific attentional effects, respectively. Importantly, both effects have been observed on the missing-item task. For example, a sequence of valent words (e.g., "hate, tease, coward…") produce greater disruption of both serial and missing item tasks than a sequence of neutral words (e.g., deer, rabbit, sheep…"; Marsh et al., 2018) and a change in voice conveying to-be remembered items (e.g., "M, J, E, **O**, B, L…") as compared to a voice-pure sequence (e.g., "M, J, E, O, B, L…") produces greater disruption of serial recall (Hughes et al., 2007; Experiment 2). As outlined earlier (see 6.1), a goal of the present study was to investigate whether the distracter familiarity effect was a specific instance of interference-by-process or attributable to a specific form of attentional capture. Results fell in line with the duplex mechanism account. That distracter melody familiarity impaired performance of a non-seriation-based task undermined interruption based on a notion that the distracter familiarity effect was an "enhanced changing-state effect" due to strongly represented order cues through top-down/schema-driven processes (e.g., Bregman, 1990), or stronger recruitment of the SMA (e.g., Lima et al., 2016). Given that results of Study IV clearly align with an attentional capture view of the distracter familiarity effect, one might attempt to speculate on the precise underpinning of such an effect.

6.7 What is the basis of the distracter familiarity attentional capture effect?

What governs the attentional capturing prowess of a familiar over an unfamiliar distracter melody? While a definitive answer to this question is not possible given the data, it is nevertheless possible to speculate. On first hearing an auditory sequence, such as a melody or song, a neural model is fashioned of the order of acoustic elements within the sequence, and this acts as a predictive model as to what will occur within the sound stream. There is an expectation created and held in long-term memory for future reference. Therefore, if expectation, based on the neural model, is violated, aspecific attentional capture will occur (e.g., Eimer et al., 1996). Although this predictive model can draw upon knowledge of a familiar melody from long-term memory, for an unfamiliar melody there is no top-down support. In this view, a well-fashioned neural model has potential to decrease disruption produced by a to-be-ignored sound. This is one explanation for why foreknowledge about, for example, a sentence, reduces the disruptive impact that sentence later has to short-term memory (Hughes & Marsh, 2019). Expectation that the neural model for a familiar melody would be particularly well fashioned could lead to supposition that familiar melody should be less disruptive than unfamiliar melody. However, in this current study the reverse is true. Therefore, predictability cannot explain the distracter familiarity effect in the same way as it can explain disruption of short-term memory for random word sequences or complex sentences (Bell et al., 2017; Hughes & Marsh, 2019).

One possibility is that mere personal significance drives the distracter familiarity effect as it does the own-name effect. One’s own name is a self-reliant stimulus that is readily activated within the working memory system such that its external presentations can be readily detected (Moray, 1959; Wood & Conway, 1995) thereby promoting withdrawal of attention from the focal task (Conway et al., 2001; Röer et al., 2013). Like one’s own name, familiar melodies are overlearned and may, therefore, also be readily activated within working memory, increasing the propensity for them to capture attention. A second, related hypothesis, is that familiar melody may be associated with an emotional response, possibly tied to activation of autobiographical memories, and this response may trigger an amygdala-driven response (Silva Pereira et al., 2011) thereby activating a distributed neural network that supports attentional responses that may not be task-relevant (cf. Bach, Furl, Barnes, & Dolan, 2015). Third, it is possible that a familiar melody, learned in concert with nursery rhyme lyrics (Thomas, 1930), leads to automatic retrieval of associated lyrics (Pring & Walker, 1994). Indeed, musical imagery is most active for melody and sung-lyrics, pitch with word-based imagery reportedly the strongest (Bailes, 2007). It is possible that enactment of lyrics or persistence of a mental image of those lyrics (via musical auditory imagery, see Chapter I, 1.7) during an unfolding familiar melody (e.g., Halpern, 1988) is attentionally capturing, diverting attention, thereby impairing focal task performance regardless of its underpinning cognitive process(es).

This argument is, however, inconsistent with findings from other studies in this empirical series. If a familiar melody is prone to automatically activate those lyrics with which it is associated (Pring & Walker, 1994) this could, in part, account for the distracter familiarity effect evidenced in Studies II-IV. However, this view is inconsistent with the lack of a distracter familiarity effect for melody production in Study I. The presence of subsumed or imagined lyrics in the familiar distracter melodies had no influence on the production (by humming) of a different familiar melody with associated lyrics from long-term-memory, as compared to the presence of an unfamiliar melody without associated lyrics. The apparent lack of automatic lyrics activation (Study I), therefore, renders this argument flawed. Questions to be answered therefore, are does predictability make familiar melody less/more interesting and thus more capable of diverting attention? And is relevance key to capturing attention? It would be interesting to see whether pre-exposure to familiar melody would reduce its disruptive effect on short-term memory since one could argue benefit of foreknowledge may be diminished by virtue of capacity to form a detailed neural model of changing-acoustic elements order within a familiar melody.

6.8 Conclusion

Study IV clearly identifies poorer task-performance in the presence of irrelevant familiar melodies (without their associated lyrics) as compared to irrelevant unfamiliar melodies (without lyrics) and quiet. This pattern of results, from a task that did not necessitate serial order processing, is in line with those from Study III that did require serial-order processing. Previous literature has suggested two possible explanations of this finding. First, a conflict of serial-order processes between processing of physical changes in a familiar to-be-ignored item, and serial order processing required for a to-be-recalled item (interference-by-process: Jones & Tremblay, 2000). Second, the potential potency of a familiar, as compared to unfamiliar, melody, to divert attention away from the focal task (Marsh et al., 2018). Crucially, as the distracter familiarity effect emerged on serial recall (Study III) and on the missing-item task (Study IV), whereby attentional capture effects are not task-process dependent (Hughes et al., 2007; Marsh et al., 2018), these findings support an attentional diversion account over an interference-by-process account of auditory distraction by familiar melody. As such, they align with the duplex mechanism account (Hughes et al., 2007).

CHAPTER VII

GENERAL OVERVIEW OF EMPIRICAL STUDIES AND THEORETICAL IMPLICATIONS

The current unique empirical series of study extends knowledge relating to the impact on music cognition from familiar and unfamiliar task-irrelevant sound comprising independent melody, sung-lyrics within a cappella song, and spoken-lyrics. Study I-II explored long-term memory retrieval and performance of known melody/lyrics, during task irrelevant sound in regard to a process-based account of auditory distraction. Moreover, the current series incorporated visual-verbal item short-term memory recall tasks (Study III-IV). This allowed opportune comparison to assess the impact of identical distracter conditions on non-music tasks requiring (Study III), and not-requiring (Study IV), serial rehearsal in order to augment the theoretical debate regarding mechanisms pertaining to short-term memory distraction. Furthermore, the familiarity of the irrelevant melody/lyrics was considered as a factor in the disruptive magnitude to task performance. In particular, the degree to which familiar background melody/lyrics well-formed motor plans generated competition with target production within the vocal-motor system, thereby exacerbating production performance of target melody/lyrics, was a key area of analysis.

To summarize, performance across Studies I to IV was more accurate in quiet conditions as compared to during all irrelevant background sound conditions. This was as expected and follows patterns previously reported from music-related tasks during irrelevant instrumental music, song, speech, or noise (e.g., Schlittmeier et al., 2008; Williamson et al., 2010), serial-recall tasks (e.g., Alley & Greene, 2008; Colle & Welsh, 1976; Jones et al., 1992; Jones & Macken, 1993; McCorkell, 2012; Pring & Walker, 1994; Salamé & Baddeley, 1982, 1989; Tremblay et al., 2001), and non-music related tasks in the presence of irrelevant speech such as mental arithmetic (Banbury & Berry, 1998; Perham et al., 2016), or comprehension (Martin et al., 1988; Perham et al., 2005; Perham & Currie, 2014). Across Studies I to III sung-lyrics were more disruptive than instrumental melody or spoken-lyrics (Alley & Greene, 2008; Iwanago & Ito, 2002; not relevant in Study IV). A crucial finding from this current series was a familiarity effect obtained for melody (familiar compared to unfamiliar) during the semantic (Study II), serial (Study III), and non-serial (Study IV) recall tasks. At first it was unclear from Study III results if this familiarity effect, whereby a familiar melody was more disruptive than an unfamiliar melody to serial recall, resulted from interference-by-process or attentional diversion. This was because it only manifested for instrumental melody, not when combined with lyrics. The fact that the familiarity effect also observed in a task thought not to require seriation (Beaman & Jones, 1998; Study IV, missing-item) signposts attentional capture as the mechanism undergirding the effect. However, some of the effects observed in the sung-lyrics conditions in Study I-II could not be explained according to the attentional capture account. This was because although potent distracters for melody/lyrics retrieval those attributes of sound contained in sung-lyrics were not disruptive for serial recall (Study III).

Irrespective of melody familiarity, the combination of melody with familiar sung-lyrics was the most potent distracter of Study I-II, as compared to instrumental melody (without associated lyrics) or spoken-lyrics (Speech). The additional component of vocal performance highlighted the interplay of the vocal-motor production mechanism for song and its vulnerability—as with speech—to the pre-attentive processing of the distracter melody/lyrics (Hughes & Marsh, 2017). Sung-lyrics were also the most potent disrupters of serial recall for Study III, but here, irrespective of melody familiarity. It appears that the speech motor planning mechanism is more vulnerable to irrelevant lyrics when sung than when spoken. This is since sung-lyrics were shown to be more effective in their ability to disrupt the vocal-motor planning for production of target item irrespective of the nature of the task, long-term memory retrieval via performance or, the more frequently adopted short-term memory recall.

Differences in auditory conditions from prior research.

These distinctive studies contrast to other theoretical approaches to short-term memory recognition/recall tasks in several ways. Routinely, task irrelevant sound is heard during short-term memory target item presentation or in a retention interval (e.g., Alley & Greene, 2008; Iwanago & Ito, 2002; McCorkell, 2012; Thompson & Yankeelov, 2012). Although for Study III-IV irrelevant auditory conditions were presented during presentation of visual target stimuli for later recall in quiet, for Study I-II concurrent vocal production of a complete different familiar melody/lyrics target was demanded <u>during</u> the irrelevant auditory conditions.

Unlike previous studies (Pring & Walker, 1994; Iwanaga & Ito, 2002) a quiet baseline condition was included for all studies. Furthermore, care was taken to maintain musical properties across all irrelevant auditory conditions. For example, consistent genre, tempo, and dynamic of irrelevant stimuli (compared to different musical excerpts/styles; Iwanaga & Ito, 2002; Perham & Vizard, 2011). A consistent female vocalist delivered sung- and spoken-lyrics conditions (compared to male and female carriers; Alley & Greene, 2008) thereby eliminating environmental origin sub-streaming (Bregman, 1990). An overview of findings according to process-based accounts of disruption to focal task by auditory distraction, and possible explanations of identified effects, now follows.

7.1 Interference-by-content *versus* interference-by-process

For Study I-II the passive, interference-by-content view was pitched against the functional, interference-by-process view (Hughes & Jones, 2005; Jones & Tremblay, 2000) in the context of familiar song melody and lyrics retrieval from long-term memory. According to the interference-by-content account, whereby disruption occurs at the item-level, familiar melodies should be no more distracting than unfamiliar melodies, regardless of the presence of lyrics. However, according to the interference-by-process account, familiar melody with or without lyrics, should produce more disruption than unfamiliar melody, with or without lyrics. This latter view was upheld.

Study I-II showed unequivocally that both familiar and unfamiliar distracters impaired retrieval. Furthermore, melody and speech, regardless of whether familiar or unfamiliar, impaired performance to a lesser extent than familiar melody and lyrics combined. That irrelevant melodies were more disruptive than spoken-lyrics when trying to produce a melody (Study I), but unfamiliar spoken-lyrics were more disruptive than melodies when trying to produce lyrics (Study II), thereby producing a dissociation, aligns to the idea that the interference-by-process concept originating in the context of the classical ISE in the context of serial recall—wherein auditory distraction is attributed to a conflict between seriation processes (Jones & Tremblay, 2000)—may be extended to the articulatory-based seriation processes involved in melody production and the processes underpinning the semantic retrieval of lyrics (Jones et al., 2004; Marsh et al., 2009). In the latter cases, the disruption is framed as being attributable to the inappropriate, irrelevant semantic information producing competition for selection mechanisms necessary for the long-term memory semantic retrieval of the target melody/lyrics (Marsh et al., 2009). The fact that planning, and production of a melody/lyrics contextually-appropriate at a general level to the irrelevant melody/lyrics, but unrelated and incompatible at the response level, is in keeping with the selection-for-action approach (Hughes & Jones, 2005). Rather than resulting from limitations to the cognitive system (Cowan, 1995; Neath, 2000) this view is embedded in the process-oriented approach whereby adaptive mechanisms are engaged to resolve competition required to select appropriate information for specific action (target retrieval) from contextually-appropriate, but response-inappropriate information for the specific action (Anderson, 2003; Marsh et al., 2009).

The fact that melody produced greater disruption than spoken-lyrics is also at odds with the interference-by-content account since, according to this view, non-speech sounds have no direct access to the phonological store (Baddeley, 1986) unless via a selective filter (Salamé & Baddeley, 1989). Similarly, that familiar melody in combination with lyrics (regardless of their familiarity) drove an additional disruption of melody recall is further inexplicable within the phonological store account. While the interference-by-process account might predict familiar melody without lyrics to be significantly more disruptive than unfamiliar melody without lyrics, which did occur in Studies II, III, and IV, but did not occur in Study I, (although familiar melody did impair recall accuracy to a greater degree than an unfamiliar melody), that familiarity when embedded with sung-lyrics was the most potent distracter coheres with the notion that these two properties make specific demands on the vocal-motor system that require inhibition to override. Most notably, this aspect of the data was inconsistent with the interference-by-content account.

Current results also relate to the supposition of embodied cognition whereby musical performance, during concurrent irrelevant musical sound, demands an interplay of musical perception, motor-planning, and action (e.g., Leman & Maes, 2014). A viewpoint based on the perceptual-gestural view (e.g., Jones et al., 2007) is that motor-systems required to produce the target performance is hampered by the presence of an irrelevant musical performance, more so when irrelevant music is familiar to the performer. The fact that for Study I-II the ability to amass and sub-vocally practice the melody/lyrics sequence patterns was disrupted to a greater extent by irrelevant sound containing lyrics adheres with the perceptual-gestural view, whereby vocal-motor sequence-planning is actively engaged to organize the to-be-remembered items into a coherent stream (Bregman, 1990). If this streaming process is reliant on the sequence of melody/lyrics, rather than the individual items (Hughes & Jones, 2005), the interplay between the irrelevant auditory input and sub-vocal gestures of to-be-performed sequences would affect performance accuracy, especially familiar items that have established representations in memory. It did.

According to the interference-by-process view, the changing-state of the distracter stimuli is reflected in the magnitude of impairment to a task, especially if tasks involve maintaining serial order (Hughes et al., 2007). Song melodies and lyrics both necessitate a set order of progression to be maintained in auditory perception and vocal production. For the current series of study, changing-state equivalence was maintained in the unfamiliar melodies and unfamiliar lyrics. Therefore, expectation would be for the increased changing-state of melody and lyrics in combination to be of greater detriment to target recall. This was demonstrated. Performance decreased in Study I-III in the presence of melody combined with lyrics, irrespective of lyric familiarity, but to a significantly greater extent with familiar lyrics in Study I when trying to produce a dissimilar yet familiar melody (lyrics not relevant in Study IV).

That a familiarity effect was noted across all studies in this current series (familiar being more disruptive than unfamiliar) is a key finding, but this effect was modulated according to auditory condition and task. In Study I-II, familiar as compared to unfamiliar melody, song, and speech, caused greater impairment to melody and lyrics recall performance. However, in Study I this was only significant for the Unfamiliar Lyrics Sound Type Group, whereas in Study II there was a familiarity effect for the Melody Group and the Familiar Lyrics Group. This effect could be argued as being derived from the familiar melody triggering the familiar lyrics that created conflict with the sub-vocal/motor planning production of the target lyrics (Pring & Walker, 1994). However, as retrieval of familiar melody from long-term memory was the only study not to show a familiarity effect for melody it would seem an implausible explanation. It could, however, suggest that interference from a lyrical, tonal, musical contour impedes the motor-system required for production of a different lyrical, tonal, musical contour irrespective of familiarity. Assuming again that retrieval of melody from long-term memory involves some representation within a memory system prior to production, the general view is that similarity between target and to-be-ignored representations should drive the disruption. However, similar problems arise for this perspective. For example, to explain why familiar melody accompanied by familiar or unfamiliar lyrics impairs melody production more than unfamiliar lyrics and unfamiliar melody combined, one would have to argue that the similarity in representation of the former to a target melody is somehow greater than for the latter. It is difficult to envisage how this could be possible: they both had the same vocal delivery and followed the same rhythmic pattern. In addition, great care was taken to match the structure, harmonic framework, tempo, and dynamic. The fact that retrieval of familiar melody from long-term memory was the only study not to show a familiarity effect for melody suggests that interference from a lyrical, tonal, musical contour impedes the motor-system required for production of a different lyrical, tonal, musical contour irrespective of familiarity.

Therefore, an alternative explanation is that the motor-plan formation necessary to produce the target melody was affected in a different way than when required to produce the lyrics. This would imply some degree of separate representation of melody and lyrics processing components within the cognitive system, such as within the musical lexicon, where direct competition would ensue between familiar to-be-performed and to-be-ignored material (Peretz & Coltheart, 2003). Interestingly, as reported earlier, there was a familiarity effect in Study I, but only for the Unfamiliar Lyrics Group, suggesting the greater incongruity of a familiar melody combined with novel words added to the disruption when trying to recall and hum the familiar target melody. So, it would seem that irrelevant melody without lyrics is not dependent on its familiarity in order to impair the retrieval and performance of a target melody (Study I).

Furthermore, both the short-term memory recall tasks (Study III-IV) also exhibited a familiarity effect for melody. Crucially, no main effect of familiarity was evident in Study III when the irrelevant sound contained lyrics, either spoken or sung. This would imply that for a short-term serial recall task, lyrics impaired performance irrespective of their familiarity, and that melody familiarity, although producing a familiarity effect when heard without lyrics, was unimportant when combined with lyrics. These results are particularly coherent with the interference-by-process view, task demands driving the effect. Although previous research has shown that in tasks that do not involve serial order recall, such as the missing-item task, irrelevant sound will produce only minor, or even no disruption to task performance as compared to quiet (Beaman & Jones, 1997; Perham et al., 2007), this was not found to be the case in Study IV of this series (missing-item task). This was an unexpected result. However, the missing-item task is known to exhibit an effect of sound (typically steady-state, as compared to changing-state, effects; LeCompte, 1995). A familiarity effect of instrumental melody was found in Study IV, mirroring Study II-III, but contrary to Study I. A possible explanation for this could be that during the unfamiliar melody condition a search for lyrics, that do not exist in long-term memory, may be triggered thereby using some cognitive capacity needed for the focal task. However, the familiar condition causes greater impairment as the familiar associated lyrics are accessed and thereby provide some competition for focal task processes. This clearly indicates an interference-by-process as opposed to a structural interference-by-content explanation (Marsh et al., 2009).

7.2 Interference-by-process *versus* attentional capture

Although the mere presence of task-irrelevant sound is clearly disruptive of short- and long-term memory, the mechanisms underpinning auditory distraction remain less clear. While disruption identified in Study I-II is consistent with the notion that familiar melody and lyrics compete for the vocal-motor apparatus required for production of the target melody/lyrics, it is possible that proponents of different theories would promote different accounts for these phenomena.

On the duplex-mechanism account (Hughes, 2014), disruption occurs due to interference-by-process—whereby pre-attentive order-encoding of auditory changes clash with serial-order cues relating to visual-verbal items—and attentional diversion, whereby sound diverts attention away from the focal task. Interference-by-process thus occurs only when serial-order is required for the focal task (Hughes, 2014). Attentional diversion, however, occurs regardless of whether the focal task requires serial-order. In Study I-II disruption to humming and speaking performance of the target melody/lyrics may have resulted from specific attentional capture driven by personal significance or emotional response to auditory stimuli (e.g., Marsh et al., 2018; Röer et al., 2013). In this approach the disruptive effects observed in Study I-II would be dissociated by the task-demands of Study III.

The key finding was that performance in Study III was disrupted significantly more by familiar distracters. But only when the sequence comprised an instrumental melody. This result clearly indicates that familiar melody produces more disruption than unfamiliar melody of short-term memory for item recall in serial order (familiarity effect). However, the basis of the familiarity effect could not be determined from this task. Crucially, the familiarity effect was also observed in the context of the missing-item task (Study IV). It was expected that familiar melody would impair serial recall (Study III), but not the missing-item task (Study IV) which, although sensitive to attentional diversion, is immune to the changing-state effect (Beaman & Jones, 1997; Elliott et al., 2016; Hughes et al., 2007). The lack of an interaction between distracter familiarity and task (serial *vs.* missing-item) in the comparative analysis suggests that this effect occurs with equal magnitude regardless of whether focal task seriation is involved. To reiterate, the surprising and critical key finding from Study III-IV was not that familiar melody distracters did produce more disruption of serial recall and missing-item task performance than unfamiliar melody distracters, but, crucially, familiar distracters were <u>not</u> more disruptive than unfamiliar distracters when combined with lyrics in song (Study III), suggesting that the disruption from lyrics overrode the disruptive impact of familiar melody.

On hearing music, humans make predictions about the impending unfolding contour thereby engaging similar cognitive mechanisms for music perception and performance (Koelsch, Vuust, & Friston, 2018; Pearce, 2018). If perception and action are linked, interaction between perceptual and motor networks will be generated during the prediction process (Monroy et al., 2019), the sensory system alerted to the most likely forthcoming sound (Morillon, 2017; Tillman, Janata, Birk, & Bharucha, 2008). This appears to be particularly true for trained musicians (Bianco, Novembre, Keller, & Scharf, 2016). The idea that forewarning of the auditory distracter, as in the case of a familiar melody/lyrics, may reduce the changing-state effect with a lessening of disruption was not evident across any of these present studies. Previous research (e.g., Hughes et al., 2013; Röer et al., 2015) has shown changing state to be attenuated by foreknowledge. Based on Röer et al. (2015), familiarity with the melody or lyrics of the distracter nursery rhyme in this series of study should have lessened the distractibility. (Although, Hughes and Marsh [2017] attribute this to the attenuation of an attentional diversion effect.) As there was no significant effect between familiar and unfamiliar melody distracters in Study I this would appear not to be so. Rather, these results support Hughes and Marsh (2017) who identified no role for predictability in the changing-state ISE. However, in the short-term memory tasks (Study III-IV), irrespective of seriation requirement, the familiarity/predictability/forewarning of the changing state, in keeping with the attentional capture idea (Cowan, 1995; Nöstl et al., 2012), did produce a familiarity effect during irrelevant familiar melody. Collectively, Study III-IV results are compatible with the two functionally different distraction types, interference of rehearsal and attentional capture, proposed by the duplex-mechanism account (Hughes et al., 2007).

Another important finding for Study I was that the onset time (OST) required to begin production of the target melody was slowest during the familiar sung-lyrics condition, suggesting the OST measure captured the temporary impedance that familiar melody added to familiar lyrics caused by capturing the articulators. Familiar Lyrics Sound Type was significant against all other Sound Types irrespective of melody familiarity. This result was not replicated when trying to speak the target lyrics in Study II or during serial-digit recall (Study III, [not relevant for Study IV]). This implies two possible explanations. First, when lyrics are decoupled from melody, there is less demand on the speech mechanism to produce the target lyrics, or, second, music is entwined to a greater degree with associated memories than lyrics, exhibiting a greater capacity to elicit disruption to the vocal-planning mechanism as irrelevant material competes with target production.

Like language, music, vocal or instrumental is not just a vehicle for delivering semantics (meaning) but also communicates emotion (whereby a lyric or/and melody can be onomatopoetic: “brilliant”, “melancholy”, or “eerie”). The emotion (prosody) of the sound shown to have greater capacity than the semantics to influence impairment on serial and non-serial-based task performance (Kattner & Ellermeier, 2018). An unfolding familiar melody has the capacity to awaken other associated schema, such as autobiographical memories related to emotions such as happiness and nostalgia (Belfi, Karlan, & Tranel, 2015; Janata, Tomic, & Rakowski, 2007), and the extent of their familiarity would, acting in competition, have had added impact on motor-sequence planning especially for retrieval of target song components (Marsh et al., 2009). These involuntary memories may be recovered directly in response to the familiar music (El Haj, Fasotti, & Allain, 2012) heard or imagined, and may have also impacted on the streaming process; music having been found more potent than other external stimuli, such as faces, to evoke autobiographical memories (Belfi et al., 2015).

That familiar melody without lyrics in Study I was not significantly more disruptive than unfamiliar melody without lyrics suggests that lyrics were not automatically activated by the familiar melody *per se* (see Pring & Walker, 1994). However, it could be speculated that humming a target melody generated the associated lyrics that potentially may have caused some added disruption, and as a consequence, increased vocal-motor action that set competition with that required for target output. Equally, the associated melody/lyrics of the irrelevant sound may have impacted in a similar way. Nevertheless, the fact that familiar melody in combination with familiar lyrics was the most potent distracter suggests that humming of a target melody may invoke the retrieval of associated lyrics (and therefore access the phonological lexicon) rather than some peripheral activation of the motor-system. These results, therefore, add weight to a process-oriented approach to attentional selectivity during retrieval from long-term memory. Although speculative, it may be considered that the visual-verbal presence of the target throughout performance may have also contributed to impairment (Farrell & Abrams, 2014). This was noted, although not recorded by the author, as a number of participants deliberately closed their eyes or turned away from the computer screen during performance to arguably, “block out” the written target rhyme title.

7.3 Instrumental *versus* vocal

The way we listen to or imagine music can produce the same neural activity (Halpern & Zatorre, 1999; King, 2006; Kraemer et al., 2005), the speech-planning mechanism identified as a medium through which melody can also be internally generated (Gathercole & Baddeley, 1993; Smith et al., 1992). Therefore, a possible explanation as to why task accuracy reduced during the familiar, as compared to the unfamiliar, melody Sound Conditions, may, in part, either be attributable to the imagined corresponding lyrics being triggered by the irrelevant melody (the involuntary use of musical imagery [Bailes, 2007] purported to impair task performance when no overt/audible lyrics present through the use of subvocalization, Pring & Walker, 1994) or, speculatively, the visual stimulus of the visual-verbal target generated the corresponding/associated melody/lyrics that may have impeded the required vocal-motor planning necessary to recall and perform the target melody/lyrics (Brandimonte, Hitch, & Bishop, 1992).

It was unambiguous that the greatest decrement to accuracy of melody recall (Study I), spoken-lyrics recall (Study II), and serial-recall (Study III) was from melody and lyrics combined within song. This is in line with studies that show vocal music to impact to a greater extent than instrumental music on tasks that require recall in order of presentation—serial recall—than sound without lyrics (e.g., Alley & Greene, 2008: Iwanaga & Ito, 2002; Marsh et al, 2009; Peretz, et al., 2009; Salamé & Baddeley, 1989; Schlittmeier & Hellbrück, 2008). This finding from the current empirical series offers to suggest this is also true for non-serial recall tasks. Furthermore, the combination of familiar melody with familiar lyrics, as opposed to unfamiliar lyrics, demonstrated the most substantial disruption in melody and lyrics recall (Study I-II). Interestingly, however, spoken-lyrics alone, familiar or unfamiliar, were the least disruptive to retrieval/performance of melody and less detrimental to sung-lyrics for Study II-III.

The idea of a modality specific ISE, whereby memory for music is impaired to a greater extent during irrelevant instrumental melody as compared to irrelevant words is in line with previous findings for short-term memory tasks (e.g., Pechmann & Mohr, 1992; Schendel & Palmer, 2007; Williamson et al., 2010). The corollary expectation, that lyrics recall would be impaired to a greater extent during irrelevant lyrics as compared to irrelevant melody, was also replicated in Study II. The finding of a familiarity effect in Study III (serial recall) runs contrary to Iwanaga and Ito (2002) and Alley and Greene (2008) who reported familiarity with song lyrics to have little effect on serial recall performance, but is in line with McCorkell (2012) who showed a familiarity effect with spoken and instrumental conditions for a serial-recall task.

Integration or independence?

Although findings from this empirical series were not conclusive, the familiarity effect findings, and the relative different effects of modality specific irrelevant sound on music and non-music tasks do suggest some degree of processing independence between music and lyrics within song. Although singing or speaking words have been shown to depend on the same pathway for verbal production (although the melodic route is different, Hébert et al., 2003), they appear to exhibit shared and distinct neural correlates (Özdemir et al., 2006). Although humming requires low-level sensorimotor operations of the singing mechanism, pitch control during humming does require analogous timing of movement (Özdemir, et al., 2006) and studies suggest similar cerebral structures are required for time perception and motor timing (Schubotz et al., 2000). Findings of increased impairment to melody production from familiar melody suggest humming, as singing, accesses the purported phonological lexicon within the modular model of music processing (Peretz & Coltheart, 2003). Finding, from largely non-musically trained participants, that familiar/unfamiliar sung-lyrics, impaired melody retrieval to a greater degree than familiar/unfamiliar spoken-lyrics, suggests some degree of independence. Furthermore, that spoken-lyrics, familiar or unfamiliar, without melody, were less disruptive than familiar melody and sung-lyrics, reinforces the separation concept of melody and lyrics within song (e.g., Bonnel et al., 2001).

Musician versus non-musician.

Results from this empirical series run contrary to previous research showing musician[60] advantage for musically related tasks (Dittinger et al., 2016; Schön et al., 2004; Thompson & Yankeelov, 2010). Entrainment practice from musical training identified as the mechanism underlying cross-domain transfer to phonological/language skill enhancement (Tierney et al., 2008; Tierney & Kraus, 2014). This practice creates a neural network's overlap in music and language domains for processing auditory-motor timing (Patel, 2011, 2014). However, across all studies in this present series no significant advantage was evidenced.

7.4 Age-related cognitive decline

Previous literature has identified differences in task performance according to age related decline in cognitive ability (e.g., Hasher & Zacks, 1988; Glinsky, 2007; Tun,

[60] 7+years musical training (Ericsson et al., 1993).

O'Kane, & Wingfield, 2002). Older people shown to be more susceptible to auditory distraction (Conway et al., 2001; Hughes et al, 2013) through a reduction in the ability to inhibit irrelevant information (Lustig, Hasher, & Zacks, 2007). Some research has shown age equivalence for tasks involving irrelevant changing-state/steady-state/deviant sounds, but lower scores for older adults during serial recall (Röer et al., 2015). If cognitive ability is influential in long-term retrieval for melody/lyrics, recall scores from older participants should decrease accordingly. However, this was not evidenced from the current empirical series. Contrary to prior research age-related cognition did not detrimentally influence recall accuracy throughout this empirical series. The ability to filter out interference from elements of irrelevant song was immune to age (Guerreiro, Murphy, & Van Gerven, 2010; Röer et al., 2015), and therefore not a factor in the creation of an ISE (Klatte, Lachmann, Schlittmeier, & Hellbrück, 2010).

7.5 Factors for consideration: Limitations

Although not part of the formal assessment process, there were issues that arose during the method that possibly impacted on overall findings. For example, scores garnered from Study I may have resulted from the requirement to vocally hum a performance of the target melody which is a much more complex activity than pitch pattern recognition. Although as humans, we have an innate ability to utilize WM and sensorimotor codes for vocal performance (Levitin, 1994) it is possible that participants varied in their ability to hum, and of greater import, to compare their vocal hummed output to the internal auditory representation of the target melody. Although humming increases the internal resonance of the vocal cords, if unaccustomed to the vocal positioning an inappropriate pitch contour may occur due to limited awareness of unfolding auditory information.

Potential hummed inaccuracies pose questions regarding the validity of the instrument of measure. Vocally untrained participants produced difficult to assess dirty musical data (Müllensiefen & Wiggins 2011; Sloboda & Parker, 1985). Assessment raters needed to infer pitch contour inaccuracies. Use of a rigorous computer-based analysis matching melodic contours by use of flexible melodic similarity algorithms (Dannenberg et al., 2007; Rocamora, Cancela, & Pardo, 2014) may have yielded different scores.

Lower overall performance scores for some participants in Study I-II (noted by the author as mainly aged 18-25) may have resulted from limited knowledge of the target resulting in incomplete melody/lyrics (less bars) being assessed, rather than the effects of distraction. Or, they may have resulted from an unwillingness to try and hum/speak if they were not confident of the complete melody/lyrics. It is also feasible that some familiar distracters may have been heard as unfamiliar. Use of a Likert scale rating with opportunities for “quite familiar” and “slightly familiar” responses may have improved familiarity assessment.

Study I-II requested that target melody/lyrics be hummed/spoken. Although not formally assessed, participants often “broke into song” as, for them, the melody/lyrics were inseparable (e.g., “Happy Birthday”), or as a means to recall subsequent lines (e.g., “When the wind blows” in “Rock-a-bye baby”). These identified pauses allowed the melody/lyrics to be used to cue the following lines. Consequently, the measure of OST in Study II was not a true indicator as it only measured the time needed to begin the first word of the lyrics. It should also be noted that the irrelevant spoken-lyrics delivery was paced to the tempo and rhythmic pattern of the melody/sung-lyrics conditions (Simmons-Stern et al., 2012), which sometimes created an intermittent conveyance[61] that lacked continuous flow with possible spasmodic non-interference.

7.6 Concluding summary

Based on the pattern of results in the thesis, retrieval from long-term memory in the presence of instrumental melody, a cappella song, or spoken-lyrics, familiar and unfamiliar, lies within the dynamic interference-by-process view of interference-by-forgetting thought to be founded within the short-term memory (Jones & Tremblay, 2000). The disruption and retrieval of the target melody/lyrics for Study I-II by task-irrelevant auditory material identified disruption to the vocal-motor planning process required to hum/speak the target melody/lyrics, due to competition for retrieval. Furthermore, as familiar melody and lyrics combined was not a potent cause of disruption in Study III (serial recall), this finding also aligns with a process driven account. An important theoretical consideration here is related to the distracter familiarity effect observed in Study II (lyric recall) that also obtained for tasks requiring serial rehearsal (Study III, serial-digit recall) and, surprisingly, those thought not to involve seriation strategies (Study IV, missing-item; Beaman, 2004). Interestingly, in Study III the familiarity effect only manifested when melody was heard separately, not when combined with lyrics, suggesting that in the serial recall task changing-state and/or attentional diversion mechanisms were operational. However, attentional

[61] For example, in natural speech: Jack and Jill went up the hill; In rhythmically paced speech: Jack – and Jill – went up – the hill.

capture was confirmed by the presence of a familiarity effect in Study IV. As attentional capture effects are not task-process dependent (Hughes et al., 2007; Marsh et al., 2018), findings from Study III-IV support an attentional diversion account of auditory distraction by familiar melody rather than an interference-by-process account.

A further theoretical consideration relates to the integration/independence of melody and lyrics within song. The fact that sung-lyrics impaired melody retrieval to a greater degree than spoken-lyrics, familiar or unfamiliar, would appear to support a degree of independence (Besson et al., 1998; Besson & Schön, 2001; Bonnel et al., 2001; Peretz et al., 1994) and that song is stored differently to verbal information in WM (Berz, 1995; Besson & Schön, 2001; Halpern & Müllensiefen, 2008; Schulze et al., 2011).

Future research.

It would be intriguing to build on these findings and explore the route for vocal-motor production/performance output via lyrics singing, as compared to humming/speaking during the same Sound Conditions. Singing requires a greater level of vocal co-ordination than humming, due to activation of melody and lyrics simultaneously. The perception and action interplay and pathways for the processing and production of lyrics during singing performance could then be cross-task compared with regard to motor-control. It would be interesting to see if results would support the clinical dissociation identified from neurological studies that show patients with left frontal lesions who can sing—but not speak—words (e.g., Brust, 2003).

Music is a potent communicator of human thoughts and emotions capable of evoking memories (Barrett et al., 2010; Belfi et al., 2015; Krumhansl & Zupnick, 2013). Finding a familiarity effect for melody in Study II-IV and increased disruption by familiar sung-lyrics may have stemmed from their potential to re-awaken personal memories irrespective of task process. This re-awakening may be helpful for people with some dementias such as AD. A question to be explored therefore, is, can familiar melody reactivate cognitive capacity? If, as Cuddy, Sikka, and Vanstone (2015, p. 223) state, “music is a prime stimulus for cognitive stimulation”, if music has the potential to affect how the brain works (Mannes, 2011), the use of music must be exploited by therapists/carers to maximise communication. Future research should now explore familiar melody/song in the association of routine procedures, possibly coupled with gesture, to support access for older adults to target information in real-life situations (Murakami, 2017) and support social interaction when spoken words become lost (Cross & Morley, 2008; Simmons-Stern et al., 2012).

References

Abdel Rahman, R., & Melinger, A. (2007). When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 604-614. doi: 10.1037/0278-7393.33.3.604

Abrams, D. A., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D. J., & Menon, V. (2013). Inter-subject synchronization of brain responses during natural music listening. *European Journal of Neuroscience, 37*(9), 1458-1469. doi: 10.1111/ejn.12173

Acheson, D. J., & MacDonald, M. C. (2009a). Twisting tongues and memories: Explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language, 60,* 329-350. http://dx.doi.org/10.1016/j.jml2008.12.002

Acheson, D. J., & MacDonald, M. C. (2009b). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135,* 50-68. http://dx.doi.org/10.1037/a0014411

Aiello, R. (1994). Music and language: parallels and contrasts. In R. Aiello & J. Sloboda (Eds.), *Musical Perceptions* (pp. 40-63). Oxford University Press. New York, NY.

Aleman, A., Nieuwenstein, M. R., Bocker, K. B. E., & de Hann, E. H. (2000). Music training and mental imagery ability. *Neuropsychologia, 38*(12), 1664-1668. doi: 10.1016/S0028-3932(00)00079-8

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*(4), 277-289. doi: 10.1007/s12144-008-9040-z

Allport, D. A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of 25 years. In D. E. Meyer & S. Komblum (Eds.), *Attention and performance XIV* (pp. 183-218). Cambridge, MA: MIT Press.

Alluri, V. (2011). Listening to music lights up the whole brain. *ScienceDaily*. Cited in Suomen Akatemia (Academy of Finland). Retrieved from www.sciencedaily.com/releases/2011/12/111205081731.htm

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage, 59*(4), 3677-3689. doi: 10.1016/j.neuroimage.2011.11.019

Alluri, V., Toiviainen, P., Lund, T. E., Wallentin, M., Vuust, P., Nandi, A. K., Ristaniem, T., & Brattico, E. (2013). From Vivaldi to Beatles and back: Predicting lateralized brain responses to music. *NeuroImage, 83,* 627-636. doi: 10.1016/j.neuroimage.2013.06.064

Ally, B. (2010). Music aids Alzheimer's patients in remembering new information. Boston University Medical Centre. (2010, May 13). *ScienceDaily*. Retrieved from www.sciencedaily.com/releases/2010/05/100512112314.htm

Alzheimer’s Disease International. (2015). Retrieved from www.alz.co.uk/worldreport2015

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49,* 415-445. doi: 10.1016/j.jml.2003.08.006

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 237-313). New York: Academic Press.

Anderson, M. C., Green, C., & McCulloch, K. C. (2000). Similarity and inhibition in long-term memory: Evidence for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1141-1159. http://dx.doi.org/10.1037/0278-7393.26.5.1141

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Handbook of perception and cognition* (2nd ed.), *Memory (*pp. 237-313). San Diego, CA: Academic Press.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological Review, 102*(1), 68-100. doi: 10.1037//0033-295X.102.1.68

Asaridou, S. S. & McQueen, J. M. (2013). Speech and music shape the listening brain: evidence for shared domain-general mechanisms. *Frontiers in Psychology, 4*(321), 1-14. doi: 10.3388/fpsyg.2013,00321

Astell, A. J., & Ellis, M. P. (2006). The social function of imitation in severe dementia. *Infant and Child Development, 15*(3), 311-319. doi: 10.1002/icd.455

Atherton, R. P., Chrobak, Q. M., Rauscher, F. H., Karst, A. T., Hanson, M. D., Steinert, S. W., & Bowe, K. L. (2018). Shared processing of language and music. *Experimental Psychology, 65*(1), 40-48. https://doi.org/10.1027/1618-3169/a000388

Ayotte, J., Peretz, I., Rousseau, I., Bard, C., & Bojanowski, M. (2000). Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain*, *123*(9), 1926-38. https://doi.org/10.1093/brain/123.9.1926

Bach, D. R., Furl, N., Barnes, G., & Dolan, R. J. (2015). Sustained magnetic responses in temporal cortex reflect instantaneous significance of approaching and receding sounds. *PLoS ONE, 10*(7), e0134060. doi: 10.1371/journal.pone.0134060

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology*, *18,* 362-365. https://doi.org/10.1080/14640746608400055

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

Baddeley, A. D. (1990). *Human Memory: Theory and practice.* Hove, UK: Erlbaum.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 11*(4), 417-423.

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

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63,* 1-29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D., Allen, R. J., & Hitch, G. J. (2010). Investigating the episodic buffer. *Psychologica Belgica*, *50*(3-4), 223-243. doi: 10.5334/pb-50-3-4-223

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129,* 126-145. doi: 10.1037/00963445.129.1.126

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation, Vol. 8* (pp. 47-89). New York: Academic Press. http://dx.doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A, 36*(2)*,* 233-252. https://doi.org/10.1080/14640748408402157

Baddeley, A. D., Thompson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behaviour, 14*(6), 575-589. doi: 10.1016/S0022-5371(75)80045-4

Baddeley, A. D., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language*, *24,* 490-502.

Bailes, F. (2002). Musical Imagery: Hearing and imagining music. ASIN: B00/ABLORQ. http://ethesis.whiterose.ac.uk/id/eprint/3452

Bailes, F. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music, 35*(4), 555-570. doi: 10.1177/0305735607077834

Bailes, F., Bishop, L., Stevens, C. J., & Dean, R. (2012). Mental imagery for musical changes in loudness. *Frontiers in Psychology, 3*(525). doi: 10.3389/fpsyg.2012.00525

Baird, A., & Samson, S. (2009). Memory for music in Alzheimer's disease: Unforgettable? *Neuropsychology Review, 19*(1), 85-101. doi: 10.1007/s11065-009-9085-2

Banbury, S. P., & Berry, D. C. (1997). Habituation and dishabituation to speech and office noise. *Journal of Experimental Psychology: Applied, 3*(3), 181-195. doi: 10.1037/1076-898X.3.3.181

Banbury, S. P., & Berry, D. C. (1998). Disruption of office-related tasks by speech and office noise. *British Journal of Psychology, 89*(3), 499-517. doi: 10.1111/j.2044-8295.1998.tb02699.x

Banbury, S. P., & Berry, D. C. (2005). Office noise and employee concentration: Identifying causes of disruption and potential improvements. *Ergonomics, 48*(1), 25-37. doi: 10.1080/00140130412331311390

Banbury, S. P., Macken, W. J., Tremblay, S., & Jones, D. M. (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors, 43*(1), 12-29. doi: 10.1518/001872001775992462

Barrett, F. S., Grimm, K. J., Robins, R. W., Wildschut, T., Sedikides, C., & Janata, P. (2010). Music-Evoked Nostalgia: Affect, memory, and personality. *American Psychological Association, 10*(3), 390-403. doi: 10.1037/a0019006

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research, 1161*, 65-78. doi: 10.1016/j.brainres.2007.05.045

Baumgartner, H. (1992). Remembrance of things past - music, autobiographical memory, and emotion. *Advances in Consumer Research, 19*, 613-620.

Beaman, C. P. (2004). The irrelevant sound phenomenon revisited: What role for working memory capacity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(5), 1106-1118. doi: 10.1037/0278-7393.30.5.1106

Beaman, C. P. (2005). Irrelevant sound effects amongst younger and older adults: Objective findings and subjective insights. *European Journal of Cognitive Psychology, 17*(2), 241-265. doi: 10.1080/09541440440000023

Beaman, C. P., Hanczakowski, M., Hodgetts, H., Marsh, J. E., & Jones, D. M. (2013). Memory as discrimination: What distraction reveals. *Memory & Cognition, 41*(8), 1238-1251. https://doi.org/10.3758/s13421-013-0327-4

Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(2), 459-471.

Beaman, C. P., & Jones, D. M. (1998). Irrelevant sound disrupts order information in free as in serial recall. *The Quarterly Journal of Experimental Psychology Section A, 51*(3), 615-636. doi: 10.1080/713755774

Beaman, C. P., Powell, K., & Rapley, E. (2015). Want to block earworms from conscious awareness? B(u)y gum! *The Quarterly Journal of Experimental Psychology, 68*(6), 1049-57. doi: 10.1080/17470218.2015.634142

Beaman, C. P., & Williams, T. I. (2010). Earworms (stuck song syndrome): Towards a natural history of intrusive thoughts. *British Journal of Psychology, 101*(4), 637-53. doi: 10.1348/000712609x479636

Belfi, A. M., Karlan, B., & Tranel, D. (2015). Music evokes vivid autobiographical memories. *Memory, 24*(7), 1-11. doi: 1080/09658211.2015.1061012

Bell, R., Buchner, A., & Mund, I. (2008). Age-related differences in irrelevant-speech effects. *Psychology and Aging, 23*(2), 377-391. doi: 10.1037/0882-7974.23.2.377

Bell, R., Dentale, S., Buchner, A., & Mayr, S. (2010). ERP correlates of the irrelevant sound effect. *Psychophysiology, 47*(6), 1182-1191. doi: 10.1111/j.1469-8986.2010.01029.x

Bell, R., Mund, I., & Buchner, A. (2011). Disruption of short-term memory by distractor speech: Does content matter? *The Quarterly Journal of Experimental Psychology, 64*(1), 146-168. https://doi.org/10.1080/17470218.2010.483769

Bell, R., Röer, J. P., Dentale, S., & Buchner, A. (2012). Habituation of the irrelevant sound effect: Evidence for an attentional theory of short-term memory disruption. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(6), 1542-1557. doi: 10.1037/a0028459

Bell, R., Röer, J. P., Lang, A.-G., & Buchner, A. (2019). Reassessing the token set size effect on serial recall: Implications for theories of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45,* 1432-1440. https://doi.org/10.1037/xlm0000658

Bell, R., Röer, J. P., Marsh., J. E., Storch, D., & Buchner, A. (2017). The effect of cognitive control on different types of auditory distraction: A preregistered study. *Experimental Psychology, 64*(5), 359-368. doi: 10.1027/1618-3169/a000372

Benjamin, A. S., & Bawa, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory and Language, 51*(2), 159-172. doi: 10.1016/j.jml.2004.04.001

Benton, A. L. (1997). The amusias. In A. L. Benton, M. Critchley, & R. A. Hanson, *Music and the brain* (pp. 378-397). Heinemann.

Berkowska, M., & Dalla Bella, S. (2009). Acquired and congenital disorders of sung performance: A review. *Advanced Cognitive Psychology, 5*, 69-83. doi: 10.2478/v10053-008-0068-2

Berlingeri, M., Bottini, G., Basilico, S., Silani, G., Zanardi, G., Sberna, M., . . . Paulesu, E. (2008). Anatomy of the episodic buffer: A voxel-based morphometry study in patients with dementia. *Behavioural Neurology, 19*(1-2), 29-34.

Berti, S., Münzer, S., Schröger, E., & Pechmann, T. (2006). Different interference effects in musicians and a control group. *Experimental Psychology, 53*(2), 111-116. doi: 10.1027/1618-3169.53.2.111

Berz, W. (1995). Working memory in music: A theoretical model. *Music Perception, 12*(3), 353-364. http://dx.doi.org/10.2307/40286188

Besson, M., & Faïta, F. (1995). An Event-Related Potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1278-1296.

Besson, M., Faïta, F., Perez, I., Bonnel, A., & Requin, J. (1998). Singing in the brain: Independence of lyrics and tunes. *Psychological Science, 9*(6), 494-498. doi: 10.1111/1467-9280.00091

Besson, M., & Schön, D. (2001). Comparison between language and music, *930*(1), 232-258. *Annals of the New York Academy of Sciences.* https://doi.org/10.1111/j.1749-6632.2001.tb05736.x

Bever, T. G., & Chiarello, R. J. (1974). Cerebral dominance in musicians and non-musicians. *Science, 185*(4150), 537-539.

Bever, T. G., & Chiarello, R. J. (2009). Cerebral dominance in musicians and non-musicians. *The Journal of Neuropsychiatry and Clinical Neurosciences, 21*(1), 94-97. https://doi.org/10.1176/appi.neuropsych.21.1.94

Bhatara A. 1., Yeung H. H. I., & Nazzi, T. I. (2000). Foreign language acquisition and melody singing. *Journal of Experimental Psychology and Human Perception Performance, 41*(2), 277-82. doi: 10.1037/a0038736

Bianco, R., Gold, B. P., Johnson, A. P., & Penhune, V. B. (2019). Music predictability and liking enhance pupil dilation and promote motor learning in non-musicians. *Scientific Reports, 9*(17060). https://doi.org/10.1038/s41598-019-53510-w

Bianco, R., Novembre, G., Keller, P. E., & Scharf, F. (2016). Syntax in action has priority over movement selection in piano playing: An ERP study. *Journal of Cognitive Neuroscience*, *28*(1), 41-54. https://doi.org/10.1162/jocn_a_00873

Blacking, J. (1973). How Musical Is Man? *University of Washington Press, Seattle.* ISBN10: 0295953381

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *The National Academy of Sciences*, *98*(20), 11818-11823. doi: 10.1073/pnas.191355898

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *National Neuroscience*, *2*(4), 382-387.

Bonnel, A. M., Faïta, F., Besson, M., & Peretz, I. (1998). Independence of lyrics and tunes in opera. Comparison of single and dual-task performances. *Manuscript submitted for publication.*

Bonnel, A. M., Faïta, F., Peretz, I., & Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs: Evidence for independent processing. *Perception & Psychophysics*, *63*(7), 1201-1213. doi: 10.3758/BF03194534

Bosch, I., Salimpoor, V. N., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience, 7*(534), 1-10.

Boulez, P. (1966). *Relevés d'apprenti.* Éditions du Seuil. Paris. ISBN: 2020019302 9782020019309

Boyle, P., & Coltheart, V. (1996). Effects of irrelevant sounds on phonological coding in reading comprehension and short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(2), 398-416. doi: 10.1080/713755630

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. M. (1992). Verbal recoding of visual stimuli impairs mental image transformations. *Memory & Cognition, 20*(4), 449-455.

Bregman, A. S. (1990). Auditory scene analysis: Hearing in complex environments. In S. McAdams & E. Bigand (Eds.), *Thinking in Sound* (pp. 10-36). Oxford. Oxford University Press.

Bregman, A. S. (2001). Auditory scene analysis. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 940-942). Oxford: Pergamon. doi: 10.1016/B0-08-043076-7/00663-X

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequence of tones. *Journal of Experimental Psychology*, *89*(2), 244-249. PMID: 5567132.

Broadbent, D. E. (1982). Task combination and the selective intake of information. *Acta Psychologica*, *50*(3), 253-290.

Broadbent, D. E., & Ladefoged, P. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, *31*, 1539.

Brodsky, W., Henik, A., Rubinstein, B., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception & Psychophysics*, *65*(4), 602-612. http://dx/doi.org/10.3758/BF03194586

Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers of Human Neuroscience*, *7*(320). doi: 10.3389/fnhum.2013.00320

Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. *Cognitive Brain Research*, *20*(3), 363-375. https://doi.org/10.1016/j.cogbrainres.2004.03.016

Brust, J. C. M. (2003). The cognitive neuroscience of music. In I. Peretz & R. Zatorre (Eds.), *Music and neurologist*: *A historical perspective* (pp. 181-191). Oxford University Press, Great Britain.

Buchner, A., Irmen, I., & Erdfelder, E. (1996). On the irrelevance of semantic information for the irrelevant speech effect. *The Quarterly Journal of Experimental Psychology Section A*, *49*(3), 765-779. doi: 10.1080/713755633

Buchner, A., Mehl, B., Rothermund, K., & Wentura, D. (2006). Artificially induced valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition*, *34*(5), 1055-62. https://doi.org/10.3758/BF03193252

Buchner, A., Steffens, M. C., Irmen, L., & Wender, K. F. (1998). Irrelevant auditory material affects counting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 48-67. http://dx.doi.org/10.1037/0278-7399.24.1.48

Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience, 20,* 762-778. doi: 10.1162/jocn.2008.20501

Buczylowska, D., & Petermann, F. (2016). Age-related differences and heterogeneity in executive functions: analysis of NAB executive functions module scores. *Archives of Clinical Neuropsychology, 31,* 254-262. doi: 10.1093/arclin/acw005

Burani, C., Vallar, G., & Bottini, G. (1991). Articulatory coding and phonological judgements on written words and pictures: The role of the phonological output buffer. *European Journal of Cognitive Psychology, 3*(4), 379-398. doi: 10.1080/09541449108406235

Burns, D. D. (1990). The generation effect: A test between single and multifactor theories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(6). doi: 10.1037/0278-7393.16.6.1060

Buschke, H. (1963). Relative retention in immediate memory determined by the missing scan method. *Nature, 200,* 1129-1130. doi: 10.1038/2001129b0

Buschke, H., & Hinrichs, J. V. (1968). Relative vulnerability of item information in short-term storage for the missing scan. *Journal of Verbal Learning and Verbal Behaviour, 7,* 1043-1048. doi: 10.1016/S0022-5371968)80065-9

Buzan, T. (1983). *Use both sides of your brain*. New York: E.P. Dutton. Inc.

Calvert, S. L., & Billingsley, R. L. (1998). Young children’s recitation and comprehension of information presented by songs. *Journal of Applied Developmental Psychology, 19,* 97-108.

Cassidy, G. G., & MacDonald, R. A. R. (2007). The effects of aggressive and relaxing popular music on the task performance of introverts and extraverts. *Psychology of Music, 35*(3), 517-538. https://doi.org/10.1177/0305735607076444

Cassidy, G. G., & MacDonald, R. A. R. (2009). The effects of music choice on task performance: a study of the impact of self-selected music and experimenter-selected music on driving game performance and experience. *Musicae Scientae, 13*(2), 357-386. doi: 10.1177/102986490901300207

Cassidy, G. G., & MacDonald, R. A. R. (2010). The effects of music on time perception and performance of a driving game. *Scandinavian Journal of Psychology, 51*(6), 455-464. doi: 10.1111/j.1467-9450.2010.00830.x

Chamorro-Premuzic, T., Swami, V., Terrado, A., & Furnham, A. (2009). The effects of background auditory interference and extraversion on creative and cognitive task performance. *International Journal of Psychological Studies, 1*(2), 18-24. ISSN 1918-7211

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America, 25*(5), 975-979.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press. ISBN: 9780262530071

Chomsky, N. (1988). *Language and problems of knowledge.* MIT Press. ISBN: 0262530708,978026253075

Chou, P. T. (2010). Attention drainage effect: How background music effects concentration in Taiwanese college students. *Journal of the Scholarship of Teaching and learning, 10*(1), 36-46.

Cleland, D. L., Davies, W. C., & Davies, T. C. (1963). Research in Reading. *The Reading Teacher, 16*(4), 224-228.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analysing skill in amnesia using perceptual learning. *Cortex, 17,* 273-278.

Cohen, N. J., & Squire, L. R. (1980b). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science, 210,* 207-210. https://doi.org/10.1126/science.7414331

Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behavior, 19*(6), 722-735. doi: 10.1016/S0022-5371(80)90403-X

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15*(1), 17-31. doi: 10.1016/S0022-5371(76)90003-7

Collina, S., Tabossi, P., & De Simone, F. (2013). Word production and the picture word interference paradigm: the role of learning. *Journal of Psycholinguist Research, 42*(5), 461-73. doi: 10.1007/s10936-012-9229-z

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology, 55,* 75-84.

Conway, A. A., Cowan, N., & Bunting, M. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review, 8*(2), 331-335. doi: 10.3758/BF03196169

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

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioural and Brain Sciences*, *24,* 87-114. doi: 10.1017/S0140525X01003922

Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics, 41*(5), 393-401. doi: 10.3758/BF03203031

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42-100. doi: 10.1016/j.cogpsych.2004.12.001

Cross, I., & Morley, I. (2008). The evolution of music: theories, definitions and the nature of the evidence. In S. Malloch & C. Trevarthen (Eds.), *Communicative Musicality: Exploring the basis of human companionship* (pp. 61-82). Oxford: Oxford University Press.

Cross, I., & Woodruff, G. E. (2009). Music as a communicative medium. In R. Botha & C. Knight (Eds.), *The Prehistory of Language, Vol. 1* (pp. 113-144). Oxford: Oxford University Press. ISBN 9780199545872

Crowder, R. G., Serafine, M. L., & Repp, B. (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory & Cognition, 18*(5), 469-476. https://doi.org/10.3758/BF03198480

Cuddy, L. L., & Cohen, A. J. (1976). Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology*, *28*(2), 255-270. doi: 10.1080/14640747608400555

Cuddy, L. L., & Duffin, J. M. (2005). Music, memory and Alzheimer's disease: Is music recognition spared in dementia, and how can it be assessed? *Medical Hypotheses, 64,* 229-235.

Cuddy, L. L., Duffin, J. M., Gill, S. S., Brown, C. L., Sikka, R., & Vanstone, A. D. (2012). Memory for melodies and lyrics in Alzheimer's disease. *Music Perception, 29*(5) 479-491. doi: 10.1525/mp.2012.29.5.479

Cuddy, L. L., Sikka, R., & Vanstone, A. (2015). Preservation of musical memory and engagement in healthy aging and Alzheimer's disease. *Annals of the New York Academy of Sciences, 1337,* 223-231. doi: 10.1111/nyas12617

Dalla Bella, S., & Berkowska, M. (2009). Singing proficiency in the Majority. *Annals of the New York Academy of Sciences, 1169,* 99-107. doi: 10.1111/j.1749-6632.2009.04558.x

Dalla Bella, S., Berkowska, M., & Sowiński, J. (2011). Disorders of pitch production in tone deafness. *Frontiers in Psychology, 2,* 164. doi: 10.3389/fpsyg.2011.00164

Dalla Bella, S., Berkowska, M., & Sowiński, J. (2015). Moving to the beat and singing are linked in humans. *Frontiers in Human Neuroscience, 9,* 663. doi: 10.3389/fnhum.2015.00663

Dalla Bella, S., Giguere, J.-F., & Peretz, I. (2007). Singing proficiency in the general population. *The Journal of the Acoustical Society of America, 121,* 1182-9.

Dalla Bella, S., Peretz, I., & Aronoff, N. (2003). Time course of a melody recognition. A gating paradigm study. *Perception & Psychophysics, 65*(7), 1019-1028.

Dalla Bella, S., Tremblay-Champoux, A., Berkowska, M., & Peretz, I. (2012). Memory disorders and vocal performance. *Annals of the New York Academy of Sciences, 1252,* 338-344. doi; 10.1111/j.1749-6632.2011.06424.x

Dannenberg, R. B., Birmingham, W. P., Pardo, B., Hu, N., Meek, C., & Tzanetakis, G. (2007). A comparative evaluation of search techniques for query-by-humming using the MUSART testbed. *Journal of the American Society for Information, Science and Technology, 58*(5), 687-701.

David, J. H., Jr. (1994). The two sides of music. *Austin Community College.* TCM 1603.7495

Defilippi, A. C. N., Garcia, R. B., & Galera, C. (2019). Irrelevant sound interference on phonological and tonal working memory in musicians and nonmusicians. *Psicologia:Reflexão e Critica, 32*(1), 2. https://doi.org/10.1186/s41155-018-0114-z

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science, 168,* 1604-1605.

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: from faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience, 28*(10), 1584-1602. doi: 10.1162/jocn_a_00997

Divin, W., Coyle, K., & James, D. T. T. (2001). The effects of irrelevant speech and articulatory suppression on the serial recall of silently presented lipread digits. *British Journal of Psychology, 92,* 593-616.

Dobbs, S., Furnham, A., & McClelland, A. (2010). The effect of background music and noise on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology, 25*(2), 307-313.

Dorsi, J., Viswanathan, N., Rosenblum, L. D., & Dias, J. W. (2018). The role of speech fidelity in the irrelevant sound effect: insights from noise-vocoded speech backgrounds. *Quarterly Journal of Experimental Psychology, 71*(10), 2152-2161. doi: 10.1177/1747021817739257

Dowling, W. J., & Harwood, D. (1986). *Music Cognition.* New York: Academic Press. ISBN-10: 1493304399

Eagan, D. E., & Chein, J. M. (2012). Overlap of phonetic features as a determinant of the between-stream phonological similarity effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(2), 473-481. doi: 10.1037/a0025368

Eimer, M., Nattkemper, D., Schröger, E., & Prinz, W. (1996). Involuntary attention. In O. Neumann & A. F. Sanders (Eds.), *Handbook of perception and action, Vol. 3* (pp. 389-446). London, UK: Academic Press.

El Haj, M., Fasotti, L., & Allain, P. (2012). The involuntary nature of music-evoked autobiographical memories in Alzheimer’s disease. *Consciousness and Cognition, 21*, 238-246. doi: 10.1016/jconcog.2011.12.005

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. doi: 10.1037/0096-1523.24.5.1406

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the “irrelevant speech effect”. *The Journal of the Acoustical Society of America, 102*(4), 2191-2199. doi: 10.1121/1/419596

Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition, 30*(3), 478-487. doi: 10.3758/BF03194948

Elliott, E. M., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal colour-word interference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(3). 654-667. doi: 10.1037/0278-7393.27.3.654

Elliott, E. M., & Cowan, N. (2005). Coherence of the irrelevant-sound effect: Individual profiles of short-term memory and susceptibility to task-irrelevant materials. *Memory & Cognition, 33*(4), 664-675. doi: 10.3758/BF03195333

Elliott, E. M., Cowan, N., & Valle-Inclan, F. (1998). The nature of cross-modal colour-word interference effects. *Perception & Psychophysics, 60*(5), 761-767. PMID: 9682602

Elliott, E. M., Hughes, R. W., Briganti, A., Joseph, T. N., Marsh, J. E., & Macken, W. J. (2016). Distraction in verbal short-term memory: Insights from developmental differences. *Journal of Memory and Language, 88,* 39-50. doi: 10.1016/j.jml.2015.12.008

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23. https://doi.org/10.1111/1467-8721.00160

Ericsson, K., Krampe, R., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychology Review, 100,* 363-406.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of target letters in a nonsearch task. *Perception & Psychophysics, 16,* 143-149. doi: 10.3758/BF03203267

Farrell, M. T., & Abrams, L. (2014). Picture-word interference reveals inhibitory effects of syllable frequency on lexical selection. *The Quarterly Journal of Experimental Psychology, 67*(3), 525-541. doi: 10.1080/17470218.2013.820763

Feng, S., & Bidelman, G. M. (2015). Music familiarity modulates mind wandering during lexical processing. Annual Meeting of the Cognitive Science Society (CogSci 2015), Pasadena, CA, July 22-25, 2015. Available at hyyp://citeweb.info/20152339756

Fitzroy, A. B., & Sanders, L. D. (2013). Musical expertise modulates early processing of syntactic violations in language. *Frontiers in Psychology, 11*(3), 603. doi: 10.3389/fpsyg.2012.00603

Fiveash, A., McArthur, G., & Thompson, W. F. (2018). Syntactic and non-syntactic sources of interference by music on language processing. *Scientific Reports 8*(1), 1-15, [17918]. doi: 10.1038/s41598-018-36076-x

Fiveash, A., & Pammer, K. (2014). Music and language: Do they draw on similar syntactic working memory resources? *Sage Journals, 42*(2), 190-209. https://doi.org/10.1177/0305735612463949

Furnham, A., Gunter, B., & Peterson, E. (1994). Television distraction and the performance of introverts and extroverts. *Applied Cognitive Psychology, 8*(7), 705-711. doi: 10.1002/acp.2350080708

Furnham, A., & Stephenson, R. (2007). Musical distracters, personality type and cognitive performance in school children. *Psychology of Music, 35*(3), 403-420. doi: 10.1177/0305735607072653

Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extroverts. *Ergonomics, 45*(3), 203-217. doi: 10.1080/00140130210121932

Furnham, A., Trew, S., & Snead, I. (1999). The distracting effects of vocal and instrumental music on the cognitive test performance of introverts and extraverts. *Pergamon, 27,* 381-392. doi: 10.1016/s0191-8869(98)00249-9

Gardiner, J. C., & Thaut, M. H. (2014). Musical mnemonics training (MMT). In M. H. Thaut & V. Hoemberg (Eds.), *Handbook of neurologic music therapy* (pp. 294-310). Oxford, UK: Oxford University Press.

Gates, A., & Bradshaw, J. C. (1977). The role of the cerebral hemispheres in music. *Brain & Language, 4*(3), 403-431.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working Memory and Language.* Hillsdale. NJ: Erlbaum.

Gillis, K. L. (2014). Working memory and music perception and production in an adult sample. *Undergraduate honours thesis 7.* Huron University College London, Canada. https://ir.lib.uwo.ca/psych_uht/7

Glaser, W. R. (1992). Picture naming. *Cognition, 42*(1-3), 61-105. https://doi.org/10.1016/0010-0277(92)90040-O

Glisky, E. L. (2007). Changes in cognitive function in human aging. Brain aging: Models, methods, and mechanisms. *Frontiers in Neuroscience. Boca Raton, F: CRC Press*. ISBN 9780849338182

Godøy, R. I., & Leman, M. (2009). *Musical Gestures: sound, movement, and meaning,* New York, NY, USA: Routledge.

Gooding, P. A., Isaac, C. L., & Mayes, A. R. (2005). Prose recall and amnesia: More implications for the episodic buffer. *Neuropsychologia, 43*(4), 583-587. https://doi.org/10.1016/j.neuropsychologia.2004.07.004

Gordon, H. W. (1981). The role of the right and left hemispheres in music perception. *The Journal of the Acoustical Society of America, 69*,103. https://doi.org/10.1121/1.386535

Götel, E., Brown, S., & Ekman, S.-L. (2000). Caregiver-assisted music events in psychogeriatric care. *Journal of Psychiatric Mental Health Nursing, 7*(2), 119-125.

Götel, E., Brown, S., & Ekman, S.-L. (2008). The influence of caregiver singing and background music on vocally expressed emotions and mood in dementia care: A qualitative analysis. *International Journal of Nursing Studies, 46*(4), 422-30. doi: 10.1016/j.injnurstu.2007.11.001

Graesser, A., & Mandler, G. (1978). Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. *Journal of Experimental Psychology*: *Human Learning and Memory*, *4*(1), 86-100. https://doi.org/10.1037/0278-7393.4.1.86

Graydon, J., & Eysenck, M. W. (1989). Distraction and cognitive performance. *European Journal of Cognitive Psychology, 1*(2), 161-179. doi: 10.1080/09541448908403078

Greene, R. L. (1991). Serial recall of two-voice lists: Implications for theories of auditory recency and suffix effects. *Memory & Cognition*, *19*(1), 72-78. doi: 10.3758/BF03198497

Gröussard, M., Viader, F., Hubert, V., Landeau, B., Abbas, A., Desgranges, B., Eustache, E., & Platel, H. (2010). Musical and verbal semantic memory: two distinct neural networks? *NeuroImage*, *49*(3), 2764-2773. doi: 10.1016/j.neuroimage.2009.10.039.inserm-00538395

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioural and Brain Sciences*, *27*(3), 377-396. doi: 10.1017/S0140525X04000093

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychology Review, 105*(4), 611-633. https://doi.org/10.1037/0033-295X.105.4.611-633

Guerreiro, M. J., Murphy, D. R., & Van Gerven, P. W. (2010). The role of sensory modality in age related distraction: A critical review and a renewed view. *Psychological Bulletin*, *136*(6), 975-1022. doi: 10.1037/a0020731

Gustavson, A., Hanneken, K., Moldysz, A., & Simon, B. (2013). The effects of music on short-term memory and physiological arousal. *Journal of Advanced Student Science, University of Wisconsin-Madison.* Available at Iowa Research Online: https://ir.uiowa.edu/honours_theses/72

Hailstone, J. C., Omar, R., & Warren, J. D. (2009). Relatively preserved knowledge of music in semantic dementia. *Journal of Neurology, Neurosurgery, and Psychiatry, 80*(7), 808-809. doi: 10.1136/jnnp.2008.153130

Halász, V., & Cunnington, R. (2012). Unconscious effects of action on perception. *Brain Sciences, 16*(2), 130-46. doi: 10.3390/brainsc2020130

Halin, N., Marsh, J. E., Hellman, A., Hellström, I., & Sörqvist, P. (2014). A shield against distraction. *Journal of Applied Research in Memory and Cognition, 3*(1), 31-36. http://dx.doi.org/10.1016/j.jarmac.2014.01.003

Hallam, S., Price, J., & Katsarou, G. (2002). The effects of background music on primary school pupils' task performance. *Educational Studies, 28*(2), 111-122. doi: 10.1080/03055690220124551

Halpern, A. R. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 343-443.

Halpern, A. R., & Bower, G. H. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*(1), 31. doi: 10.2307/1422658

Halpern, A., & Müllensiefen, D. (2008). Effects of timbre and tempo change on memory for music. *The Quarterly Journal of Experimental Psychology, 61*(9), 1371-1384. doi: 10.1080/17470210701508038

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex, 9*(7), 697-704. doi: 10.1093/cercor/9.7.697

Hamzelou, J. (2010). Music and lyrics: How the brain splits songs. *Article in New Scientist*, 09.03.2010.

Hanczakowski, M., Beaman, C. P., & Jones, D. M. (2017). When distraction benefits memory through semantic similarity. *Journal of Memory and Language, 94,* 61-74. http://dx.doi.org/10.1016/j.jml.2016.11.005

Hanley, J. R. (1997). Does articulatory suppression remove the irrelevant speech effect? *Memory, 5*(3), 423-31. doi: 10.1080/741941394

Hara, M. (2011). Music in Dementia Care: Increased understanding through mixed research methods. *Music and Arts in Action, 3*(2), 34-58.

Hasher, L., & Weeks, J. C. (2014). The disruptive – and beneficial – effects of distraction on older adults' cognitive performance. *Frontiers in Psychology, 5*(133). doi: 10.3389/fpsyg.2014.00133

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation, 22,* 193-255. doi: 10.1016/S0079-7421(08)60041-9

Heathcote, J. (2009). *Memories are made of this: Reminiscence activities for person-centred care.* London: Alzheimer's Society.

Hébert, S., & Peretz, I. (1997). Recognition of music in long term memory: Are melodic and temporal patterns equal partners? *Memory & Cognition, 25*(4), 518-533. doi: 10.3758/BF03201127

Hébert, S., & Peretz, I. (2001). Are text and tune of familiar songs separable by brain damage? *Brain and Cognition, 46,* 169-175. doi: 10.1016/S0278-2626(01)80058-0

Hébert, S., Peretz, I., & Gagnon, L. (1995). Perceiving the tonal ending of tune excerpts: the roles of pre-existing representation and musical expertise. *Canadian Journal of Experimental Psychology, 49*(2), 193-209.

Hébert, S., Racette, A., Gagnon, L., & Peretz, I. (2003). Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain, 126*(8), 1838-1850. doi: 10.1093/brain/awg186

Henson, R. A. (1977). Neurological aspects of musical experience. In M. Critchley & R. A. Henson (Eds.), *Music and the brain: Studies in the neurology of music* (pp. 3-21). London: Heinemann medical.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience, 15*(5), 673-682. doi: 10.1162/089892903322307393

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*(5), 339-402

Hilliard, O. M., & Tolin, P. (1979). Effect of familiarity with background music on performance of simple and difficult reading comprehension tasks. *Perceptual and Motor Skills, 49*(3), 713-714. doi: 10.2466/pms.1979.49.3.713

Ho, Y. C., Cheung, M. C., & Chan, A. S. (2003). Music training improves verbal but not visual memory cross-sectional and longitudinal explorations in children. *Neuropsychology, 17*(3), 439-450. doi: 10.1037/0894-4105.17.3.439

Hommel, B. (2010). Grounding attention in action control: The intentional control of selection. In B. J. Bruya (Ed.), *Effortless attention: A new perception in the cognitive science of attention and action* (pp. 121-140). Cambridge, MA: MIT Press. http://dx.doi.org/10.7551/mitpress/9780262013840.003.0006

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioural and Brain Sciences, 24*, 849-937. https://doi.org/10.3389/fpsyg.2015.01318

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCH Journal, 3*(1), 30-41. doi: 10.1002/pchj.44

Hughes, R. W., Chamberland, C., Tremblay, S., & Jones, D. M. (2016). Perceptual-motor determinants of auditory-verbal serial short-term memory. *Journal of*

Memory and Language, 90(Supplement C), 126-146. doi: 10.1016/j.jml.2016.04.006

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance, 39*(2), 539-553. doi: 10.1037/a0029064

Hughes, R. W., & Jones, D. M. (2001). The intrusiveness of sound: Laboratory findings and their implications for noise abatement. *Noise & Health, 4*(13), 51-70.

Hughes, R. W., & Jones, D. M. (2003). A negative order-repetition priming effect: Inhibition of order in unattended auditory sequences? *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 199-218. doi: 10.1037/0096-153.29.1.199

Hughes, R. W., & Jones, D. M. (2005). The impact of order incongruence between a task-irrelevant auditory sequence and a task-relevant visual sequence. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 316-327. doi: 10.1037/0096-1523.31.2.316

Hughes, R. W., & Marsh, J. E. (2017). The functional determinants of short-term memory: Evidence from perceptual-motor interference in verbal serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(4), 537-551. doi: 10.1037/xlm0000325

Hughes, R. W., & Marsh, J. E. (2019). When is forewarned forearmed? Predicting auditory distraction in short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication. https://doi.org/10.1037/xlm0000736

Hughes, R. W., Marsh, J. E., & Jones, D. M. (2009). Perceptual-gestural (mis)mapping in serial short-term memory: The impact of talker variability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1411-25. doi: 10.1037/a0017008

Hughes, R. W., Marsh, J. E., & Jones, D. M. (2011). Role of serial order in the impact of talker variability on short-term memory: Testing a perceptual organization-based account. *Memory & Cognition, 39*(8), 1435-1447. doi: 10.3758/s13421-011-0116-x

Hughes, R. W., Tremblay, S., & Jones, D. M. (2005). Disruption by speech of serial short-term memory: The role of changing-state vowels. *Psychonomic Bulletin and Review, 12*(5), 886-890. doi: 10.3758/BF03196781

Hughes, R. W., Vachon, F., & Jones, D. M. (2005). Auditory attentional capture during serial recall: Violations at encoding of an algorithm-based neural model? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(4), 736-749. doi: 10.1037/0278-7393.31.4.736

Hughes, R. W., Vachon, F., Hurlstone, M. J., Marsh, J. E., Macken, W. J., & Jones, D. M. (2011). *Disruption of cognitive performance by sound: Differentiating two forms of auditory distraction.* Paper presented at the 10th international congress on noise as a public health problem, London, UK.

Hughes, R. W., Vachon, F., & Jones, D. M. (2005). Auditory attentional capture during serial recall: Violations at encoding of an algorithm-based neural model? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(4), 736-749. doi: 10.1037/0278-7393.31.4.736

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 1050-1056. doi: 10.1037/0278-7393.33.6.1050

Humphreys, M. S., & Schwartz, R. M. (1971). Within category generalization in a missing scan paradigm. *Journal of Verbal Learning and Verbal Behaviour, 10,* 694-710.

Hussain, G., Thompson, W. E., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception, 20*(2), 151-71. doi: 10.1525/mp.2002.20.2.151

Hutchins, S., & Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology: General, 141,* 76-97. doi: 10.1037/a0025064

Hutchins, S., Roquet, C., & Peretz, I. (2012). The vocal generosity effect: How bad can your singing be? *Music Perception, 30*(2), 147-159. doi:10.1525/MP.2012.30.2.147

Hygge, S., Boman, E., & Enmarker, I. (2003). The effects of road traffic noise and meaningful irrelevant speech on different memory systems. *Scandinavian Journal of Psychology, 44*(1), 13-21. doi: 10.1111/1467-9450.00316

Ibbotson, N. R., & Morton, J. (1981). Rhythm and dominance. *Cognition, 9*(2), 125-138. doi: 10.1016/0010-0277(81)90008-1

Iwanaga, M., & Ito, T. (2002). Disturbance effect of music on processing of verbal and special memories. *Perceptual Motor Skills, 94*, 1251-1258. doi: 10.2466/pms.2002.94.3c.1251

Jakobson, J., Lewycky, S., Kilgour, A., & Stoesz, B. (2008). Memory for verbal and visual material in highly trained musicians. *Music Perception, 26*(1), 41-55 doi: 10.1525/mp.2008.26.1.41

Jakubowski, K., Finkl, S., Stewart, L., & Müllensiefen, D. (2016). Dissecting an earworm; Melodic features of song popularity predict involuntary musical imagery. *Psychology of Aesthetics, Creativity, and the Arts, 11*(2), 122-135. ISSN 1931-3896 (Article). doi.org/10.1037/aca0000090

Janata, P. (1995). ERP measure assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience, 7*(2), 153-164. doi: 10.1162/jocn.1955.7.2.153

Janata, P., Birk, J., Van Horn, J., Leman, M., Tillmann, B., & Bharucha, J. (2002). The cortical topography of tonal structures underlying western music. *Science, 298*(5601), 2167-70. http://hdl.handle.net/1854/LU-159558

Janata, P., Tomic, S. T., & Rakowski, S. K. (2007). Characterization of music-evoked autobiographical memories. *Memory*, *15*(8), 845-860. doi: 10.1080/09658210701734593

Jeffries, K. J., Fritz, J. B., & Braun, A. R. (2003). Words in melody: An H(2)150 PET study of brain activation during singing and speaking. *Neuroreport, 14*(5), 749-754. doi: 10.1097/01.wnr.00000066198.94941.a4

Jellison, J. A. (1976). Accuracy of temporal order recall for verbal and song digit-spans presented to right and left ears. *Journal of Music Therapy, 13*(3) 114-129.

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntactic processing in children. *NeuroImage, 47*(2), 735-744. doi: 10.1016/jneuroimage.2009.04.090

Jerison, H. J. (1959). Effects of noise on human performance. *Journal of Applied Psychology, 43*(2), 96-101.

Johnson, J., Chang, C., Brambati, S. M., Migliaccio, R., Gorno-Tempini, M. L., Miller, B. L., & Janata, P. (2011). Music recognition in frontotemporal lobar degeneration and Alzheimer's disease. *Cognitive Behavioural Neurology, 24*(2), 74-84. doi: 10.1097/WNN. 0b013e31821de326

Jones, D. M. (1993). Objects, streams and threads of auditory attention. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness and control*: *A tribute to Donald Broadbent* (pp. 87-104). New York: Clarendon Press.

Jones, D. M. (1994). Disruption of memory for lip-read lists by irrelevant speech: Further support for the changing-state hypothesis. *The Quarterly Journal of Experimental Psychology, Section A, 47*(1), 143-160. doi: 10.1348/000712699161314

Jones, D. M. (1999). The cognitive psychology of auditory distraction: The 1997 BPS Broadbent Lecture. *British Journal of Psychology, 90*(2), 167-187. doi: 10.1348/000712699161314

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., & Macken, W. J. (1999). Organisational 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*(2), 464-473. doi: 10.1037/0278-7393.25.2.464

Jones, D. M., Beaman, C. P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 209-238). London: Lawrence Erlbaum Associates.

Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 1008-1018. http://dx.doi.org/10.1037/0278-7393.21.4.1008

Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language, 54*(2), 265-281. doi:10.1016/j.jml.2005.10.006

Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology, 60*(4), 497-504. doi: 10.1080/17470210601147598

Jones, D. M., Hughes, R. W., & Macken, W. J. (2010). Auditory distraction and serial memory: The avoidable and the ineluctable. *Noise and Health, 12*(49), 201-209. doi: 10.4103/1463-1741.70497

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*(2), 369-381.

doi: 10.1037/0278-7393.19.2.369

Jones, D. M., & Macken, W. J. (1995a). Auditory babble and cognitive efficiency - role of number of voices and their location. *Journal of Experimental Psychology: Applied 1*(3), 216-226. doi: 10.1037/1076-898X.1.3.216

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*(2), 192-200. doi: 10.3758/BF03197221

Jones, D. M., & Macken, W. J. (1995c). Phonological similarity in the irrelevant speech effect: Within- or between-stream similarity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 103-115. doi: 10.1037/0278- 7393.21.1.103

Jones, D. M., Macken, W. J., & Harries, C. (1997). Disruption of short-term recognition memory for tones: Streaming or interference? *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A*(2), 337-357. doi: 10/1080/713755707

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*(3), 318-328. doi: 10.3758/BF03208264

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 656-674. doi: 10.1037/0278-7393.30.3.656

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 44*(4), 645-669. doi: 10.1080/14640749208401304

Jones, D. M., Marsh, J. E., & Hughes, R. W. (2012). Retrieval from memory: Vulnerable or inviolable? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(4), 905-922. doi: 10.1037/a0026781

Jones, D. M., Miles, C., & Page, J. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory? *Applied Cognitive Psychology, 4*(2), 89-108. doi: 10.1002/acp.2350040203

Jones, D. M., Saint-Aubin, J., & Tremblay, S. (1999). Modulation of the irrelevant sound effect by organizational factors: Further evidence from streaming by location. *The Quarterly Journal of Experimental Psychology Section A, 52*(3), 545-554. doi: 10.1080/713755832

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review, 7*(3), 550-558. doi: 10.3758/BF03214370

Jones, G. (2012). Why chunking should be considered as an explanation for developmental change before short-term memory capacity and processing speeds. *Frontiers of Psychology, 3,* 167. doi: 10.3389/fpsyg.2012.00767

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychology Review, 96,* 459-91.

Jordan, C. (2017). *Exploring a possible tonal loop in musicians and non-musicians and the relationship between musical expertise and cognitive aging.* (Unpublished Doctoral thesis.) Retrieved from Edinburgh Research Archive. URI http://hal.handle.net/1842/31077

Joseph, T. N., Hughes, R. W., Sörqvist, P., & Marsh, J. E. (2018). Differences in auditory distraction between adults and children: A duplex-mechanism approach. *Journal of Cognition, 1*(1), 1-11. [13]. https://doi.org/10.5334/joc.15

Kahneman, D., & Triesman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). Orlando, FL: Academic Press.

Kämpfe, J., Sedlmeier, P., & Renkewitz, F. (2010). The impact of background music on adult listeners: A meta-analysis. *Psychology of Music.* Advance online publication. doi: 10.1177/0305735610376261

Kattner, F., & Ellermeier, W. (2018). Emotional prosody of task-irrelevant speech interferes with the retention of serial order. *Journal of Experimental Psychology: Human Perception and Performance, 44*(8), 1303-1312. http://dx.doi.org/10.1037/xhp0000537

Kattner, F., & Ellermeier, W. (2020). Distraction at the cocktail party: Attenuation of the irrelevant speech effect after a training of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance 46*(1). doi: 10.1057/xhp0000695

Kellaris, J. J. (2001). Identifying properties of tunes that get “stuck in your head.” In *Proceedings of the Society for Consumer Psychology, Winter Conference* (pp. 66-67). Scottsdale, AZ: American Psychological Society.

King, A. J. (2006). Auditory neuroscience: activating the cortex without sound.

Current Biology, 16(11), 410-411. doi: 10.1006/nimg.2001.0832

Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short-term memory does not involve the “working memory” of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General, 112*(2), 240-264. http://dx.doi.org/10.1037/0096-3445.112.2.240

Klatte, M., & Hellbrück, J. (1993). Der “irrelevant speech effect”. Wirkungen von hintergrundschall auf das arbeitsgedächtnis. Zeitschrift Für Lärmbekämpfung, *40*(4), 91-98.

Klatte, M., Kilcher, H., & Hellbrück, J. (1995). Wirkungen der zeitlichen Struktur von Hintergrundschall auf das arbeitsgedcähtnis und ihre Theoretischen und Praktischen Implikationen (The effects of temporal structure of background noise on working memory: Theories and applications). *Zeitschrift für, Lärmbekämpfung, 40*, 91-98.

Klatte, M., Lachmann, T., Schlittmeier, S., & Hellbrück, J. (2010). The irrelevant sound effect in short-term memory: Is there developmental change? *European Journal of Cognitive Psychology, 22*(8), 1168-1191. doi: 10.1080/09541440903378250

Klatte, M., Lee, N., & Hellbrück, J. (2002). Effects of irrelevant speech and articulatory suppression on serial recall of heard and read materials. *Psychologische Beiträgc, 44*, 166-186.

Klatte, M., Meis, M., Sukowski, H., & Schick, A. (2007). Effects of irrelevant speech and traffic noise on speech perception and cognitive performance in elementary school children. *Noise and Health, 9*(36), 64-74. doi: 10.4103/1463-1741.36982

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology, 15*(2), 207-212. doi: 10.1016/j.conb.2005.03.005

Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences, 14*(3), 131-137. doi: 10.1016/j.tics.2010.01.002

Koelsch, S. (2011). Towards a neural basis of music perception – a review and updated model. *Frontiers in Psychology, 2*(6), 110. doi: 10.3389/fpsyg.2011.00110

Koelsch, S. (2013). *Brain and music.* John Wiley & Sons. ISBN: 0470683406,9780470683408

Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical “language-network” serves the processing of music, *NeuroImage, 17*(2), 956-966. doi:10.1016/S1053-8119(02)91154-7

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005b). Interaction between syntax processing in language and in music: An ERP Study. *Journal of Cognitive Neuroscience, 17*, 1565-1577. doi: 10.1162/089892905774597290

Koelsch, S., Schulze, F. T., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: an fMRI study. *NeuroImage, 25*(4), 1068-76. doi: 10.1016/j.neuroimage.2004.12.050

Koelsch, S., Vuust, P., & Friston, K. (2018). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences 23*(1), 63-77. https://doi.org/10.1016/j.tics.2018.10.006

Konecni, V. (1982). Social interaction and musical preference. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 497-516). New York, NY: Academic Press.

Körner, U., Röer J. P., Buchner, A., & Bell, R. (2017). Working memory capacity is equally unrelated to auditory distraction by changing-state and deviant sounds. *Journal of Memory and Language, 96,* 122-137. doi: 10.1016/j.jml.2017.05.005

Körner, U., Röer, J. P., Buchner, A., & Bell, R. (2018). Time of presentation affects auditory distraction: Changing-state and deviant sounds disrupt similar working memory processes. *Quarterly Journal of Experimental Psychology, 72*(3), 457-471. doi: 10.1177/1747021818758239

Kosslyn, S. M. (2005). Mental images and the brain. *Cognitive Neuropsychology, 22*(3-4), 333-347. https://doi.org/10.1080/02643290442000130

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*(9), 635-642.

Kozlov, M. D., Hughes, R. W., & Jones, D. M. (2012). Gummed-Up Memory: Chewing Gum Impairs Short-Term Recall. *Quarterly Journal of Experimental Psychology, 65*(3), 501-513. doi: 10.1080/17470218.2011.629054

Kraemer, D. J. M., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Sound of silence activates auditory cortex. *Nature, 434*(7030), 158. doi: 10.1038/434158a

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*(1), 159-179. doi: 10.1037/0033-2909.126.1.259

Krumhansl, C. L., & Zupnik, J. A. (2013). Cascading reminiscence bumps in popular music. *Psychological Science, 24*(10), 2057-2068. doi: 10.1177/0956797613486486

Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and language syntax interact in Broca’s area: An fMRI study. *PLoS ONE, 10*(11), e0141069. doi: 10.1371/journal.pone.0141069

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205. doi: 10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition, 11*(5), 539-550. http://dx.doi.org/10.3758/BF03196991

Larrouy-Maestri, P., & Morsomme, D. (2014). Criteria and tools for objectively analyzing the vocal accuracy of popular song. *Logopaedics Phoniatrics Vocology, 39,* 11-18. doi: 10.3109/14015439.2012.696139

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*(3), 145-157. doi: 10.1080/096582100387579

LeCompte, D. C. (1994). Extending the irrelevant speech effect beyond serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1396-1408. doi: 10.1037/0278-7393.20.6.1396

LeCompte, D. C. (1995). An irrelevant speech effect with repeated and continuous background speech. *Psychometric Bulletin and Review, 2,* 391-397. https//doi.org/10.3758/BF03210979

LeCompte, D. C. (1996). Irrelevant speech, serial rehearsal and temporal distinctiveness: A new approach to the irrelevant speech effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1154-1165. doi: 10.1037/0278-7393.22.5.1154

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*(2), 472-483. doi: 10.1037/0278-7393.23.2.472

LeCompte, D. C., & Shaibe, D. M. (1997). On the irrelevance of phonological similarity to the irrelevant speech effect. *The Quarterly Journal of Experimental Psychology Section A, 50*(1), 100-118. doi:10.1080/713755679

Lehmann, M., & Hasselhorn, M. (2007). Variable memory strategy use in children's adaptive intra-task learning behaviour: Developmental changes and working memory influences in free recall. *Child Development, 78*(4), 1068-1082. doi: 10.1111/j.1467-8624.2007.01053.x

Leman, M. (2007). *Embodied Music Cognition and Mediation Technology.* Cambridge, MA: MIT Press.

Leman, M., & Maes, P. J. (2014). Music perception and embodied music cognition. In L. Shapiro (Ed.), The *Routledge handbook of embodied cognition* (pp. 81-89). New York, NY: Routledge.

Leman, M., & Maes, P. J. (2015). The role of embodiment in the perception of music. *Empirical Musicology Review, 9*(3-4), 236-246. doi: 10.18061/emr.v913-4.4498

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics, 56*(4), 414-423. doi: 10.3758/BF03206733

Levitin, D. J. (2006). *This is your Brain on Music: Understanding a human obsession.* London: Atlantic Books. ISBN: 978-0-85789-514-1

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in “language” areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, *20*(4), 2142-2152. doi.org/10.1016/j.neuroimage.2003.08.016

Lewandowsky, S., & Murdock, B. B., Jr. (1989). Memory for serial order. *Psychological Review*, *96*(1), 25-57. doi: 10.1037/0033-295X.96.1.25

Liikkanen, L. A. (2009). How the mind is easily hooked on musical imagery. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg & P.-S. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music, ESCOM 2009* (pp. 271-275). Retrieved from https://jyx.jyu.fi/dspace/bitstream/handle/123456789/20890/urn_nbn_fi_jyu-2009411274. pdf?sequence=1.

Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends in Neurosciences, 39*(8), 527-542. doi: 10.1016/j.tins.2016.06.003

Ljungberg, J. K., Parmentier, F. B., Jones, D. M., Marsja, E., & Neely, G. (2014). ‘What’s in a name?’ ‘No more than when it’s mine own’. Evidence from auditory oddball distraction. *Acta Psychologica (Amst), 150,* 161-6. doi: 10.1016/j.actpsy.2014.05.009

Locke, J., & Fehr, F. (1972). Subvocalization of heard or seen words prior to spoken or written recall. *The American Journal of Psychology, 8*(1), 63-68.

Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: Recent developments in a "new view". In D. S. Gorfein & C. M. MacLeod (Eds.), *The place of inhibition in cognition* (pp. 145-162). Washington, DC: American Psychological Association.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109,* 35-54. http://dx.doi.org/10.1037/0033-295X.109.1.35

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner ear. Single or double agency? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 436-448.

Macken, W. J., Mosdell, N., & Jones, D. M. (1999). Explaining the irrelevant-sound effect: Temporal distinctiveness or changing state? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(3), 810-814. doi: 10.1037/0278-7393.25.3.810

Macken, W. J., Phelps, F. G., & Jones, D. M. (2009). What causes auditory distraction? *Psychonomic Bulletin & Review, 16*(1), 139-144. doi: 10.3758/PBR.16.1.139

Macken, W. J., Taylor, J. C., & Jones, D. M. (2014). Language and short-term memory: the role of perceptual-motor affordance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1257-1270. http://dx.doi.org/10.1037/a0036845

Macken, W. J., Taylor, J. C., Kozlov, M. D., Hughes, R. W., & Jones, D. M. (2016). Memory as embodiment: the case of modality and serial short-term memory. *Cognition, 155,* 113-124. http://dx.doi.org/10.1016/j.cognition2016.06.013

Macken, W. J., Tremblay, S., Houghton, R. J., Nicholls, A. P., & Jones, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 43-51. doi: 10.1037/0096-1523.29.1.43

Maes, P.-J., & Leman, M. (2013). The influence of body movements on children's perception of music with an ambiguous expressive character. *PloS One, 8*(1), e54682.

Maidment, D. W., & Macken, W. J. (2012). The ineluctable modality of the audible: Perceptual determinants of auditory verbal short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 989-997. http://dx.doi.org/10.1037/a0027884

Maidment, D. W., Macken, B., & Jones, D. M. (2013). Modalities of memory: Is reading lips like hearing voices? *Cognition*, *129*(3), 471-493. http://dx.doi.org/10.1016/j.cognition.2013.08.017

Mannes, E. (2011). *The Power of Music: Pioneering discoveries in the new science of song.* Walker & Company, New York. ISBN: 978-0-8027-79-829-1

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than non-musicians: Behavioural and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19*(9), 1453-1463. doi: 10.1162/jocn.2007.19.9.1453

Marsh, J. E., Beaman, C. P., Hughes, R. W., & Jones, D. M. (2012). Inhibitory control in memory: Evidence for negative priming in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1377-1388. doi: 10.1037/a0027849

Marsh, J. E., Campbell, T. A., Vachon, F., Taylor, P. J., & Hughes, R. W. (2019). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, and Psychophysics*. Advance online publication. PMID 31290133. doi: 10.3758/s13414-019-01800-w

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008). Auditory distraction in semantic memory: A process-based approach. *Journal of Memory and Language, 58*(3), 682-700. doi: 10.1016/j.jml.2007.05.002

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference-by-process, not content, determines semantic auditory distraction. *Cognition, 110*(1), 23-38. doi: 10.1016/j.cognition.2008.08.003

Marsh, J. E., Hughes, R. W., Sörqvist, P., Beaman, P., & Jones, D. M. (2015a). Erroneous and veridical recall are not two sides of the same coin: Evidence from semantic distraction in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1728-1740. doi: 10.1037/xlm0000121

Marsh, J. E., & Jones, D. M. (2010). Cross modal distraction by background speech: What role for meaning? *Noise & Health, 12*(49), 210-216. doi: 10.4103/1463-1741.70499

Marsh, J. E., Ljung, R., Jahncke, H., MacCutcheon, D., Pausch, F., Ball, I. J., & Vachon, F. (2018). Why are background telephone conversations distracting? *Journal of Experimental Psychology: Applied, 24*(2), 222-235. doi: 10.1037/xap0000170

Marsh, J. E., Perham, N., Sörqvist, P., & Jones, D. M. (2014). Boundaries of semantic distraction: Dominance and lexicality act at retrieval. *Memory & Cognition, 42*(8), 1285-301. doi: 10.3758/s13421-014-0438-6

Marsh, J. E., Röer, J., Bell, R., & Buchner, A. (2014). Predictability and distraction: Does the neural model represent post-categorical features? *PsyCh Journal*, *3*(1), 58-71. doi: 10.1002/pchj.50

Marsh, J. E., Sörqvist, P., Hodgetts, H. M., & Beaman, C. P. (2015a). Distraction control processes in free recall: Benefits and costs to performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(1), 118-133. doi: 10.1037/a0037779

Marsh, J. E., Sörqvist, P., & Hughes, R. W. (2015b). Dynamic cognitive control of irrelevant sound: Increased task engagement attenuates semantic auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance, 45*, 1462-1474. https://dx.doi.org/10.1037/xhp0000060

Marsh, J. E., Vachon, F., & Jones, D. M. (2008). When does between-sequence phonological similarity promote irrelevant sound disruption? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 243-248. http://dx.doi.org/10.1037/0278-7393.34.1.243

Marsh, J. E., Vachon, F., & Sörqvist, P. (2017). Increased distractibility in schizotype: Independence of individual differences in working memory capacity? *The Quarterly Journal of Experimental Psychology, 70*(3), 565-578. doi: 10.1080/17470218.2016.1172094

Marsh, J. E., Yang, J., Qualter, P., Richardson, C., Perham, N., Vachon, F., & Hughes, R. W. (2018). Post-categorical auditory distraction in serial short-term memory: Insights from increased task load and task type. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(6), 882-897. doi: 10.1037/xlm0000492

Marsja, E. (2011). Attention Capture: Studying the distracting effect of one´s own name. (Dissertation.) Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-46607

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language, 27*(4), 382-398. http://dx.doi.org/10.1016/0749-596X(88)90063-0

Maybery, M. T., Parmentier, F. B. R., & Jones, D. M. (2002). Grouping of list items reflected in the timing of recall: Implications for models of serial verbal memory. *Journal of Memory and Language, 47*(3), 360-385. http://dx.doi.org/10.1016/S0749-596X(02)00014-1

McCorkell, N. (2012). Is speech special? The impact of familiarity with irrelevant background sound on working memory. Available at https://e-space.mmu.ac.uk

McCullough Campbell, S., & Margulis, E. H. (2015). Catching an earworm through movement. *Journal of New Music Research, 44*(4). doi: 10.1080/09298215.2015.1084331

McDermott, J., & Hauser, M. (2005). The origins of music innateness, uniqueness, and evolution. *Music Perception, 23*(1), 29-59. ISSN 0730-7829

McElrea, H., & Standing, L. (1992). Fast music causes fast drinking. *Perceptual and Motor Skills, 75*(2), 362. doi: 10.2466/PMS.75.5.362-362

McGeoch, J. A. (1942). *The Psychology of Human Learning.* New York: Longmans, Green.

Mella, N., Fagot, D., & de Ribaupierre, A. (2016). Dispersion in cognitive functioning: age differences over the lifespan. *Journal of Clinical Experimental Neuropsychology, 38,* 111-126. doi: 10.1080/13803395.2015.1089979

Melby-Lervåg, M., & Hulme, C. (2010). Serial and free recall in children can be improved by training: Evidence for the importance of phonological and semantic representations in immediate memory tasks. *Psychological Science, 21*(11), 1694-1700. http://dx.doi.org/10.1177/0956797610385355

Meng, Z., Lan, Z., Yan, G., Marsh, J. E., & Liversedge, S. P. (2020). *Task demands modulate the effect of speech on text processing*. Submitted manuscript.

Mensink, J. G., & Raaijmakers, J. G. W. (1988). A model of interference and forgetting. *Psychological Review, 95*(4), 434-455.

Meyer, L. (1956). *Emotion and Meaning in Music.* University of Chicago Press. Chicago, IL.

Midorikawa, A., Kawamura, M., & Kezuka, M. (2003). Musical alexia for rhythm notation: a discrepancy between pitch and rhythm. *Neurocase, 9,* 232-38.

Miles, C., Jones, D. M., & Madden, C. A. (1991). Locus of the irrelevant speech effect in short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(3), 578-584. http://dx.doi.org/10.1037/0278-7393.17.3.578

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review, 63*(2), 81-97. doi: 10.1037/h0043158

Miozzo, M., & Caramazza, A. (2003). When more is less: a counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General, 132*(2), 228-52. doi: 10.1037/0096-3445.132.2.228

Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *NeuroImage, 38*(2), 331-345, doi: 10.1016/j.neuroimage.2007.07.034

Mithen, S. (2005). *The Singing Neanderthals: The origins of Music, Language, Mind, and Body.* London: Weidenfeld and Nicolson, ISBN: 0-297-64317-7

Molino, J. (2000). Towards an evolutionary theory of music and language. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music* (pp. 165-176). MIT Press, Cambridge, MA.

Monroy, C. D., Meyer, M., Schröer, L., Gerson, S. A. & Hunnius, S. (2019). The infant motor predicts actions based on visual statistical learning. *Neuroimage 185*, 947-954. doi: 10.1016/j.neuroimage.2017.12.016

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *The Quarterly Journal of Experimental Psychology, 1*(1), 56-60. doi: 10.1080/17470215908416289

Morillon, B. (2017). Predicting when a sound will occur relies on the brain's motor system. *Science Daily, Proceedings of the National Academy of Sciences*. Retrieved from www.sciencedaily.com/releases/2017/10/171005141732.htm

Morris, N., & Jones, D. M. (1990). Habituation to irrelevant speech: Effects on a visual short-term memory task. *Perception & Psychophysics, 47*(3), 291-297. doi: 10.3758/BF03205003

Morris, N., Jones, D. M., & Quayle, A. (1989). Memory disruption by background speech and singing. In E. D. Megaw (Ed.), *Contemporary Ergonomics* (pp. 494-499). London: Taylor and Frances.

Morrison, A. B., Rosenbaum, G. M., Fair, D., & Chein, J. M. (2016). Variation in strategy use across measures of verbal working memory. *Memory & Cognition, 44*(6), 922-936. doi: 10.3758/s13421-016-0608-9

Morton, J. (1990). The development of event memory. *The Psychologist, 1,* 3-10.

Moussard, A., Bigand, E., Belleville, S., & Peretz, I. (2012). Music as an aid to learn new verbal information in Alzheimer's Disease. *Music Perception, 29*(5), 521-531. http://dx.doi.org/10.1525/mp.2012.29.5.521

Mowsesian, R., & Heyer, M. R. (1973). The effect of music as a distraction on task-taking performance. *Measurement and Evaluation in Guidance, 6*(2), 104-110. https://doi.org/10.1080/00256307.1973.12022580

Mueller, G. E., & Pilzecker, A. (1900). Experimentalle Beitrage zur Lehre com Gedachtnis. *Zeitschrift fur Psychologie, 1,* 1-300.

Müllensiefen, D., & Wiggins, G. A. (2011). Sloboda and Parker's recall paradigm for melodic memory: a new, computational perspective. In I. Deliege & J. Davidson (Eds.), *Music and the mind*: *Essays in honour of John Sloboda* (pp. 161-188). Oxford: Oxford University press. ISBN: 9780199581566

Mulligan, N. W. (2002). The generation effect: Dissociation enhanced item memory and disrupted order memory. *Memory & Cognition, 30*(6), 850-861. https://doi.org/10.3758/BF03195771

Münzer, S., Berti, S., & Pechman, T. (2002). Encoding of timbre, speech, and tones: Musicians vs. non-musicians. *Psychologische Beltrage, 44*(2), 187-202. https://www.researchgate.net/publication/225/88659

Münzer, S., & Pechmann, T. (2000). Concurrent processing of tonal and verbal materials in working memory: Do musicians differ from non-musicians? In E. Schröger, A. Mecklinger & A. Friederici (Eds.), *Working on working memory* (pp. 79-96). Leipzig, Germany: Leipziger Universitätsverlag.

Murakami, B. (2017). Music as a mnemonic device for verbal recall in healthy older adults. *Open Access Theses*, 666. Retrieved from https://scholarlyrepository.miami.edu/oa_theses/666

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology, 21,* 263-276.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, *78*(4, Pt 1), 679-684. doi: 10.1037/h0026641

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*(3), 251-269. https://doi.org/10.3758/BF03213879

Nairne, J. S., Reigler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 702-709. http://dx.doi.org/10.1037/0278-7393.17.4.702

Nantais, K. M., & Schellenberg, E. G. (1999). The Mozart effect: An artefact of preference. *Psychological Science, 10*(4), 370-373. https://doi.org/10.1111/1467-9280.00170

Neath, I. (2000). Modelling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review, 7*(3), 403-423. doi: 10.3758/BF03214356

Neath, I., & Nairne, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review*, *2*(4), 429-441. doi: 10.3758/BF03210981

Neath, I., & Surprenant, A. M. (2001). The irrelevant sound effect is not always the same as the irrelevant speech effect. In H. L. Roediger III, J. S. Nairne, I. Neath & A. M. Surprenant (Eds.), *Science conference series. The nature of remembering: Essays in honour of Robert G. Crowder* (pp. 247-265). Washington, DC, US: American Psychological Association. doi: 10.1037/10394-013

Neely, C. B., & LeCompte, D. C. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory & Cognition, 27*(1), 37-44. doi: 10.3758/BF03201211

Neely, J. H., & Kahan, T. A. (2001). Is semantic activation automatic? A critical re-evaluation. In H. L. Roediger III, J. S. Nairne, I. Neath & A. M. Surprenant (Eds.), *Science conference series. The nature of remembering: Essays in honour of Robert G. Crowder* (pp. 69-93). American Psychological Association. https://doi.org/10.1037/10394-005

Neisser, U. (1967). Cognitive Psychology. New York, NY: Appleton Century-Crofts.

Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J., & Ullman, M. T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguist Research, 30*(3), 339-364. doi: 10.1023/A:101049911

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 361-394). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Neumann, O. (1996). Theories of attention. In O. Neumann & A. F. Sanders (Eds.), *Handbook of perception and action, Vol. 3* (pp. 389-446). London: Academic Press.

Nicholls, A. P., & Jones, D. M. (2002). The sandwich effect reassessed: Effects of streaming, distraction, and modality. *Memory & Cognition, 30*(1), 81-88. https://doi.org/10.3758/BF03195267

Nittono, H. (1997). Background instrumental music and serial recall. *Perceptual and Motor Skills 84*, 1307-1313. doi: 10.2466/pms.1997.84.3c.1307

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*(5), 1093-1105. https://doi.org/10.1037/0278-7393.30.5.1093

North, A. C., & Hargreaves, D. J. (1997). Music and consumer behavior. In A. C. North & D. J. Hargreaves (Eds.), *The social psychology of music* (pp. 268-89). Oxford: Oxford University Press. ISBN: 9780198523833

North, A. C., & Hargreaves, D. J. (2008). *The social and applied psychology of music.* Oxford: Oxford University Press. ISBN-10: 0198567421

North, A. C., Hargreaves, D. J., & O'Neill, S. A. (2000). The importance of music to adolescents. *British Journal of Educational Psychology, 70*(2), 225-272. http://dx.doi.org/10.1348/000709900158083

Nöstl, A., Marsh, J. E., & Sörqvist, P. (2012). Expectations modulate the magnitude of attentional capture by auditory events. *PLoS ONE, 7*(11), e48569. doi: 10.1371/journal.pone.0048569

Nunes-Silva, M., & Haase, V. G. (2013). Amusias and modularity of musical cognitive processing. *Psychology & Neuroscience, 6*(1), 45-56. doi: 10.3922/j.psns.2013.1.08

Oberauer, K., & Lange, E. B. (2008). Interference in working memory: Distinguishing similarity-based confusion, feature overwriting, and feature mitigation. *Journal of Memory and Language*, *58*(3), 730-745.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A. Cowan, N., . . . Ward, G. (2018). Benchmarks for models of short-term and working memory, *Psychological Bulletin, 144*(9), 885-958. http://dx.doi.org/10.1037/bul0000153

Oswald, C. J. P., Tremblay, S., & Jones, D. M. (2000). Disruption of comprehension by the meaning of irrelevant sound. *Memory, 8,* 345-350. doi: 10.1080/09658210050117762

Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *Neuroimage*, *33*(2), 628-635.

Page, M., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761-781. doi: 10.1037/0033-295X.105.4.761-781

Palisson, J., Roussel-Baclet, C., Maillet, D., Belin, C., Ankri, J., & Narme, P. (2015). Music enhances verbal episodic memory in Alzheimer's disease. *Journal of Clinical Experimental Neuropsychology, 37*(5), 503-17. doi: 10.1080/13803395.2015.1026802

Parmentier, F. B., & Andrés, P. (2010). The involuntary capture of attention by sound; novelty and postnovelty distraction in young and older adults. *Experimental Psychology, 57,* 68-76. doi: 10.1027/1618-3169/a000009

Parmentier, F. B. R., Elford, G., Escera, C., Andrés, P., & San Miguel, I. (2008). The cognitive locus of distraction by acoustic novelty in the cross-modal oddball task. *Cognition*, *106*(1), 408-432. doi: 10.1016/j.cognition.2007.03.008

Parmentier, F. B. R., Pacheco-Unguetti, A. P., & Valero, S. (2018). Food words distract the hungry: Evidence of involuntary semantic processing of task-irrelevant but biologically-relevant unexpected auditory words. *PLoS ONE, 13*(1), e0190644. https://doi.org/10.1371/journal.pone.0190644

Patel, A. D. (1998). Syntactic processing in language and music: Different cognitive operations, similar neural resources? *Music Perception, 16*(1), 27-42. http://dx.doi.org/10.2307/40285775

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674-681. doi: 10.1038/nn1082

Patel, A. D. (2008). *Music, Language, and the Brain.* New York: Oxford University Press. ISBN: 978-0-19-975530-1

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology, 2*, 142. doi: 10.3389/fpsyg.2011.00142

Patel, A. D. (2014). Can non-linguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis: *Hearing Research, 308,* 98-108. doi: 10.10161/j.heares.2013.08.011

Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences, 1423*(1), 378-395. https://doi.org/10.1111/nyas.13654

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition, 20*(3), 314-320. https://doi.org/10.3758/BF03199668

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113*(4), 1185-1205. doi: 10.1093/brain/113.4.1185

Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology, 10*(1), 21-56. doi: 10.1080/02643299308253455

Peretz, I. (1996). Can we lose memories for music? The case of music agnosia in a non-musician. *Journal of Cognitive Neuroscience, 8*(6), 481-496. doi: 10.1162/jocn.1996.8.6.481

Peretz, I. (2001). Brain specialization for music. New evidence from congenital amusia. *Annals of the New York Academy of Sciences*, *930,* 153-65. doi: 10.1111/j.1749-6632.2001.tb05731x

Peretz, I. (2003). Brain specialization for music: New evidence from congenital amusia. In I. Peretz & R. J. Zatorre (Eds.), *The Cognitive Neuroscience of Music* (pp. 192-203). Oxford: Oxford University Press.

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition, 100*(1), 1-32. Epub, 2006, Feb 20.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron 33*(2), 185-91. doi: 10.1016/80896-6273(01)00580-3

Peretz, I., & Babai, M. (1992). The role of contour and intervals in the recognition of melody parts: Evidence from cerebral asymmetries in musicians. *Neuropsychologia, 30*(3), 277-292.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences, 999,* 58-75. doi: 10.1196/annals.1284.006

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 688-691. doi:10.1038/nn1083

Peretz, I., Gagnon, L., Hébert, S., & Macoir, J. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Perception, 21*(3), 373-390. doi: 10.1525/mp.2004.21.3.373

Peretz, I., Gaudreau, D., & Bonnel, A. M. (1998). Exposure effects on music preference and recognition. *Memory & Cognition, 26,* 884-902. https://doi.org/10.3758/BF03201171

Peretz, I., Gosselin, S., Belin, P., Zatorre, R. J., Plailly, J., & Tillmann, B. (2009). Music lexical networks: The cortical organisation of music recognition. *Annals of the New York Academy of Sciences, 1169,* 256-265. doi: 10.1111/j.1749-6632.2009.04557.x

Peretz, I., Kolinsk, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain 117*(6), 1283-1301.

Peretz, I., Radeau, M., & Arguin, M. (2004b). Two-way interactions between music and language: evidence from priming recognition of tunes and lyrics in familiar songs. *Memory & Cognition, 32*(1), 142-152. http//dx.doi.org/10.3758/BF03195827

Peretz, I., Vuvan, D., Lagrois, M.-E., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 370*(1664), 20140090. doi:10.1098/rstb.2014.0090

Peretz, I., & Zatorre, R. (2003). *The cognitive neuroscience of music.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525202.001.0001

Peretz, I., & Zatorre, R. J. (2005). Brain organisation for music processing. *Annual Review of Psychology, 56,* 89-114. doi: 10.1146/annurev.psych.56.091103.070225

Perham, N., Banbury, S., & Jones, D. M. (2007). Reduction in auditory distraction by retrieval strategy. *Memory, 15*(4), 465-473. doi: 10.1080/09658210701288244

Perham, N., & Currie, H. (2014). Does listening to preferred music improve reading comprehension performance? *Applied Cognitive Psychology*, *28*(2), 279-284. doi: 10/1002/acp.2994

Perham, N., Hodgetts, H., & Banbury, S. (2013). Mental arithmetic and non-speech office noise: an exploration of interference-by-content. *Noise and Health*, *15*(62), 73-78. doi:10.4103/1463-1741.107160

Perham, N., Marsh, J. E., Clarkson, M., Lawrence, R., & Sörqvist, P. (2016). Distraction of mental arithmetic by background speech. *Experimental Psychology, 63*(3), 141-149. doi: 10.1027/1618-3169/a000314

Perham, N., & Sykora, M. (2012). Disliked music can be better for performance than liked music. *Applied Cognitive Psychology*, *26*, 550-554. doi: 10.1002/acp.2826

Perham, N., & Vizard, J. (2011). Can preference for background music mediate the irrelevant sound effect? *Applied Cognitive Psychology*, *25*(4), 625-631. doi: 10.1002/acp.1731

Perruchet, P., & Poulin-Charronnat, B. (2013). Challenging prior evidence for a shared syntactic processor for language and music. *Psychonomic Bulletin & Review, 20*(2), 310-7. doi: 10.3758/s13423-012-0344-5

Peynircioğlu, Z. F., Rabinovitz, B. E., & Thompson, J. L. W. (2008). Memory and metamemory for songs: the relative effectiveness of titles, lyrics, and melodies as cues for each other. *Sage Journals, 36*(1), 47-61. doi: 10.1177/0305735607079722

Peynircioğlu, Z. F., Tekcan, A. I., Wagner, J. L., Baxter, T. L., & Shaffer, S. D. (1998). Name or hum that tune: feeling of knowing for music. *Memory & Cognition, 26*(6), 1131-1137. http://doi.org/10.3758/BF03201190

Pfordresher, P. Q. (2006). Coordination of perceptual action in music performance. *Advances in Cognitive Psychology, 2*(2-3), 183-198.

Pfordresher, P. Q., Brown, S., Meier, K. M., Belyk, M., & Liotti, M. (2010). Imprecise singing is widespread. *The Journal of the Acoustical Society of America, 128*(4), 2182-2190.

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Mozart effect–Shmozart effect: A meta-analysis. *Intelligence, 38*(3), 314-323. https://doi.org/10.1016/j.intell.2010.03.001

Plailly, J., Tillman, B., & Royet, J. R. (2007). The feeling of familiarity of music and odours: the same neural signature? *Cerebral Cortex, 17*(11*),* 2650-2658. doi: 10.1093/cercor/bhl173

Platel, H. (2005). Functional neuroimaging of semantic and episodic musical memory. *Annual New York Academy of Sciences, 1060*, 136-147. doi: 10.1196/annals.1360.010

Platel, H., Baron, J. C., Desgranges, B., Bernard, F., & Eustache, F. (2003). Semantic and episodic memory of music are subserved by distinct neural networks. *NeuroImage*, *20*(1), 244-256.

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., . . . Eustache, F. (1997). The structural components of music perception: a functional anatomical study. *Brain*, *120*(2), 229-243.

Pool, M. M. (2002). The impact of background radio and television on high school students' homework performance. *Journal of Communication*, *53*(1), 74-87. doi: 10.1093/joc/53.1.74

Pool, M. M., Koostra, C. M., & van der Voort, T. H. A. (2003). Distraction effects of background television on homework performance. Research article. https://doi/10.1260/0957456021499225

Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition, 94*(3), B67-B78.

Prickett, C. A. (1974). *A comparison of two rhythmic patterns as aids to digit recall.* Unpublished Master's Thesis, Florida State University, Tallahassee.

Prickett, C. A., & Moore, R. S. (1991). The use of music to aid memory in Alzheimer's patients. *Journal of Music Therapy, 28*(2), 101-110. doi: 10.1093/jmt/28.2.101

Pring, L., & Walker, J. (1994). The Effects of unvocalized music on short-term memory. *Current Psychology, 13*(2), 165-171. doi: 10.1007/BF02686799

Prinz, W., Beisert, M., & Herwig, A. (2013). *Action Science: Foundations of an emerging discipline*. Cambridge MA: MIT Press. doi: 10.755/mitpress/978026201855.003.0001

Racette, A., & Peretz, I. (2007). Learning lyrics: to sing or not to sing? *Memory & Cognition, 35*(2). 242-253. https://doi.org/10.3758/BF03193445

Raffman, D. (1993). *Language, Music, and Mind.* Cambridge, MA: MIT Press.

Rainey, D. W., & Larsen, J. D. (2002). The effect of familiar melodies on initial learning and long-term memory for unconnected text. *Music Perception, 20*, 173-186. doi: 10.1525/MP.2002.20.2.173

Rathbone, C. J., O'Connor, A. R., & Moulin, C. J. A. (2017). The tracks of my years: Personal significance contributes to the reminiscence bump. *Memory & Cognition, 45*(1), 137-150. doi: 10.3758/s1342

Rathus, S. A. (2012). *Psychology: Concept and Connections. Brief version.* Cengage Learning.

Rauscher, F. H., Shaw G. L., & Ky, K. N. (1993). Music and spatial task performance. *Nature, 365,* 611. doi: 10.1038/365611a0

Rauscher, F. H., Shaw G. L., & Ky, K. N. (1995). Listening to Mozart enhances spatial-temporal reasoning: Towards a neurophysiological basis. *Neuroscience Letters, 185,* 44-47.

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading.* Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., & Pollatsek, A. (1994). *The Psychology of Reading.* London: Prentice Hall. ISBN: 9780805818727

Reisberg, D., Rappaport, I., O'Shaughnessy, M. (1984). Limits of working memory: the digit-digit-span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(2), 203-21. doi: 10.1037/0278-7393.10.2.203

Reisberg, D., Smith, J. D., Baxter, D., & Sonenshine, M. (1989). "Enacted" auditory images are ambiguous; "pure" auditory images are not. *Quarterly Journal of Experimental Psychology, 41*(3-A), 619-641. https://doi.org/10.1080/14640748908402385

Reisberg, D., Wilson, M., & Smith, J. D. (1991). Auditory imagery and inner speech. In R. Logie & M. Denis (Eds.), *Mental Images in Human Cognition* (pp. 59-81). Elsevier, Amsterdam.

Risset, J. C. (1991). Speech and music combined: An overview. In J. Sundberg, L. Nord & R. Carlson (Eds.), *Music, Language, Speech and Brain* (pp. 368-379). Cambridge, England: Cambridge University Press. https://doi.org/10.1007/978-1-349-12670-5_35

Roberts, L. (1986). Modality and suffix effects in memory for melodic and harmonic musical materials. *Cognitive Psychology, 18*(2), 123-157.

Rocamora, M., Cancela, P., & Pardo, A. (2014). Query by humming: Automatically building the database from music recordings. *Pattern Recognition Letters, 36*, 272-280. doi: 10.1016/j.patrec.2013.04.006

Roediger, H. L., Gallo, D. A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory, 10*, 319-332. doi: 10.1080/09658210224000144

Röer, J. P., Bell, R., & Buchner, A. (2013). Self-relevance increases the irrelevant speech effect: Attentional disruption by one's own name. *Journal of Cognitive Psychology*, *25*(8), 925-931. doi: 10.1080/20445911.2013.828063

Röer, J. P., Bell, R., & Buchner, A. (2015). Specific foreknowledge reduces auditory distraction by irrelevant speech. *Journal of Psychology: Human Perception and Performance, 41*(3), 692-702. doi: 10.1037/xhp0000028

Röer, J. P., Bell, R., Körner, U., & Buchner, A. (2019). A semantic mismatch effect on serial recall: Evidence for interlexical processing of irrelevant speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 515-525. doi: 10.1037/xlm0000596

Röer, J. P., Körner, U., Buchner, A., & Bell, R. (2017). Attentional capture by taboo words: A functional view of auditory distraction. *Emotion, 17*(4), 740-750. doi: 10.1037/emo0000274

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211-227. https://doi.org/10.1037/0096-3445.126.3.211

Rosenbaum, D. A. (2009). *Human Motor Control.* Cambridge, UK: Academic Press. ISBN: 9780123742261

Sacks, O. (2006). The power of music. *Brain, 129*(10), 2528-2532. https://doi.org/10.1093/brain/awl234

Sacks, O. (2007). *Musicophilia: Tales of music and the brain.* Picador. ISBN: 978-0-330-47113-8

Saito, S. (1997). When articulatory suppression does not suppress the activity of the phonological loop. *British Journal of Psychology, 88*(4), 565-578. http://dx.doi.org/10.1111/j.2044-8295.1997.tb02658.x

Saito, S. (1998). Phonological loop and intermittent activity: A whistle task as articulatory suppression. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 52*(1), 18-24. http://dx.doi.org/10.1037/h0087275

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada T., Iwata, N. K., & Nieleen, M. (1999). Neural Representation of a rhythm depends on its interval ratio. *Journal of Neuroscience, 19*(22), 10074-10081.

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behaviour, 21*(2), 150-164. doi: 10.1016/S0022-5371(82)90521-7

Salamé, P., & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society, 24*(4), 263-265. doi: 10.3758/BF03330135

Salamé, P., & Baddeley, A. D. (1987). Noise, unattended speech and short-term memory. *Ergonomics, 30*(8), 1185-1194. doi: 10.1080/0014938708966007

Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology, Section A, 41*(1), 107-122. doi: 10.1080/14640748908402355

Salimpoor, V. N., & Zatorre, R. J. (2013). Neural interactions that give rise to musical pleasure. *Psychology of Aesthetics, Creativity, and the Arts, 7*(1), 62-75. doi: 10.1037/a0031819

Sammler, D., Baird, A., Valabrégue, R., Clement, S., Dupont, S., Berlin, P., & Samson, S. (2010). The relationship of lyrics and tunes in the processing of unfamiliar songs: a functional magnetic resonance adaption study. *Journal of Neuroscience, 30*, 3572-3578. doi: 10.1523/JNEUROSCI.2751-09.2010

Sammler, D., Novembre, G., Koelsch, S., & Keller, P. E. (2013). Syntax in a pianist's hand: ERP signatures of 'embodied' syntax processing in music. *Cortex, 49*(5), 1325-1339. https://doi.org/10.1016/j.cortex.2012.06.007

Samson, S., & Peretz, I. (2005). Effects of prior exposure on music liking and recognition in patients with temporal lobe lesions. *Annals of the New York Academy of Sciences, 1060*, 419-28.

Samson, S., & Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: Evidence for dual-encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 793-804. doi: 10.1037/0278-7393.17.4.793

Sänger, J., Muller, V., & Lindenberger, U. (2012). Intre and interbrain synchronization and network properties when playing guitars in duets. *Frontiers in Human Neuroscience.* doi: 10.3389/fnhum.2012.00312

Schaal, N., Javadi, A.-H., Halpern, A. R., Bettina, P., & Banissy, M. J. (2015). Right parietal cortex mediates recognition memory for melodies. *European Journal of Neuroscience, 42*(1), 1660-1666. doi: 10.1111/ejn.12943

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*(8), 511-514. doi: 10.1111/j.0956-7976.2004.00711.x

Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science, 14*(6), 317-320. https://doi.org/10.1111/j.0963-7214.2005.00389.x

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology, 98*(2), 457-468. http://dx.doi.org/10.1037/0022-0663.98.2.457

Schellenberg, E. G., & Hallam, S. (2005). Music listening and cognitive abilities in 10 and 11-year-olds: The Blur effect. *Annals of the New York Academy of Sciences, 1060,* 202-209. doi:10.1196/annals.1360.013

Schellenberg, S., & Moore, R. S. (1985). The effect of tonal-rhythmic context on short term memory of rhythmic and melodic sequences. *Bulletin of the Council for Research in Music Education, 85,* 207-217.

Schendel, Z. A., & Palmer, C. (2007). Suppression effects on musical and verbal memory. *Memory & Cognition, 35*(4), 640-650. https://doi.org/10.3758/BF03193302

Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995). In viro evidence of structural brain asymmetry in musicians. *Science, 267*(5198), 699-701. https://doi.org/10.1126/science.7839149

Schlaug, G., Marchina, S., & Norton, A. (2008). From singing to speaking: Why singing may lead to recovery of expressive language function in patients with Broca's asphasia. *Music Perception, 25*(4), 315-323. doi: 10.1525/MP.2008.25.4.315

Schlaug, G., Marchina, S., Norton, A., & Wan, C. Y. (2010). From singing to speaking: facilitating recovery from nonfluent aphasia. *Future Neurology, 5*(5), 657-665.

Schlittmeier, S. J., & Hellbrück, J. (2009). Background music as noise abatement in open-plan offices: A laboratory study on performance effects and subjective preferences. *Applied Cognitive Psychology, 23*(5), 684-697. doi: 10.1002/acp.1498

Schlittmeier, S. J., Hellbrück, J., & Klatte, M. (2008). Does irrelevant music cause an irrelevant sound effect for auditory items? *European Journal of Cognitive Psychology, 20*(2), 252-271. doi: 10.1080/09541440701427838

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésane, C., Anton, J.-L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage, 51*(5), 450-461. doi: 10.1016/j.neuroimage.2010.02.023

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341-349. doi: 10.1111/1469-8986.00172x

Schubotz, R. I. (2007). Prediction of external events with our motor systems: towards a new framework. *Trends in Cognitive Sciences, 11*(5), 211-8. doi: 10.1016/j.tics.2007.02.006

Schubotz, R. I., Friederici, A. D., von Cramon, D. Y. (2000). Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *NeuroImage*, *11*(1), 1-12. https://doi.10.1006/nimg.1999.0514

Schulze, K., & Koelsch, S. (2012). Working memory for Speech and Music. *Annals of the New York Academy of Sciences*, 1252, 229-36. doi: 10.1111/j.1749-6632.2012.06447.x

Schulze, K., Mueller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *European Journal of Neuroscience, 33*(1), 189-196. doi: 10.111/j.1460-9568.2010.07470.x. Epub 2010 Nov14.

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in non-musicians and musicians. *Human Brain Mapping, 32*(5), 771-783. doi: 10.1002/hbm.21060

Schutz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: action-induced modulations of perception. *Trends in Cognitive Sciences, 11*(8), 349-55. doi: 10.1016/j.tics,2007.06.005

Schwartz, K. D., & Fouts, G. T. (2003). Music preferences, personality style, and developmental issues of adolescents. *Journal of Youth and Adolescents, 32*(3), 205-213. doi: 10.1023/A:1022547520656

ScienceDaily, (2011). *Suomen Akatemia (Academy of Finland).* Retrieved from www.sciencedaily.com/release/2011/12/111205081731.htm

Segalowitz, S. J. (1983). *Two sides of the brain. Brain lateralization explored.* Englewood Cliffs, New Jersey: Prentice Hall.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science, 275*(5306), 1599-1603. https://doi.org/10.1126/science.275.5306.1599

Serafine, M. J., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for songs. *Cognition*, *16*(3), 285-303. PMID: 6541107. doi: 10.101610012-0277(84)90031-3

Serafine, M. J., Davidson, J., Crowder, R. G., & Repp, B. H. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language, 25*, 123-135. doi: 10.1016/0749-596X(86)90025-2

Sheffert, S. M., Pisoni, D. B., Fellows, J. M., & Remez, R. E. (2002). Learning to recognize talkers from natural, sinewave, and reversed speech samples. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1447-169. doi: 10.1037/0096-1523.28.6.1447

Shih, Y., Huang, R., & Chiang, H. (2009). Correlation between work concentration levels and background music: A pilot study. *IOS Content Library.* http://content.iospress.com/articles/work/wor00880.

Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences USA*, *110*(1), 70-75. doi: 10.1073/pnas.1209023110

Silva Pereira, C., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattics, E. (2011). Music and emotions in the brain: Familiarity matters. *PloS One, 6*(11), https://doi.org/10.1371/journal.pone.0027241

Silverman, M. J. (2007). The effects of paired pitch, rhythm, and speech on working memory as measured by sequential digit recall. *Journal of Music Therapy*, *44*(4), 415-427. PMID: 17997628

Silverman, M. J. (2010). The effect of pitch, rhythm, and familiarity on working memory and anxiety as measured by digit recall performance. *Journal of Music Therapy, 47*(1), 70-83.

Simmons-Stern, N. R., Budson, A. E., & Ally, B. A. (2010). Music as a memory enhancer in patients with Alzheimer's disease. *Neuropsychologia, 48*(10), 3164-3167. doi: 10.1016/j.neuropsychologia.290.04.033

Simmons-Stern, N. R., Deason, R. G., Brandler, B. J., Frustcae, B. S., O’Connor, M. K., Ally, B. A., & Budson, A. E. (2012). Music-based memory enhancement in Alzheimer’s Disease: promise and limitations. *Neuropsychologia, 50*(12), 3295-3303.

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review, 16*(2), 374-81. doi: 10.3758/16.2.374

Sloboda, J. A., & Parker, D. (1985). Immediate recall of melodies. In P. I. C. Howell & R. West (Eds.), *Musical Structure and Cognition* (pp. 143-167). London: Academic Press.

Smith, J. D., Reisberg, D., & Wilson, M. (1992). Subvocalization and auditory imagery: Interactions between the inner ear and inner voice. In D. Reisberg (Ed.), *Auditory Imagery* (pp. 95-120). Hillsdale, NJ: Erlbaum.

Smith, J. D., Wilson, M., & Reisberg, D. (1995). The Role of Subvocalization in Auditory Imagery. *Neuropsychologia, 33*(11), 1433-1454. doi: 10,1016/0028-3932(95)0074-D

Sokolov, E. N. (1963). Higher nervous functions; the orienting reflex. *Annual Review of Physiology, 25,* 545-580. doi: 10.1146/annurev.ph.25.030163.002553

Sörqvist, P. (2010a). High working memory capacity attenuates the deviation effect but not the changing-state effect: Further support for the duplex-mechanism account of auditory distraction. *Memory & Cognition, 38*(5), 651-658. doi: 10.3758/MC.38.5.651

Sörqvist, P., & Marsh, J. E. (2015). How concentration shields against distraction. *Current Directions in Psychological Science, 24*, 267-272. doi:10.1177/0963721415577356

Sörqvist, P., Nöstl, A., & Halin, N. (2012). Disruption of writing processes by the semanticity of background speech. *Scandinavian Journal of Psychology, 53*(2), 97-102. doi: 10.1111/j.1467-9450.2011.00936.x

Steinke, W. R., Cuddy, L. L., & Jakobson, L. S. (2001). Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. *Cognitive Neuropsychology, 18*(5), 411-37. doi: 10.1080/02643290125702

Sternberg, S. (1969). The discovery of processing stages: Extension of Donder's method. *Acta Psychologica, 30,* 276-315.

Sternberg, S., Wright, C. E., Knoll, R. L., & Monsell, S. (1980). Motor programs in rapid speech: Additional evidence. In R. A. Cole (Ed.), *Perception and Production of Fluent Speech* (pp. 507-534). Hillsdale, NJ: Erlbaum.

Stroop, J. R. (1935). Interference in serial verbal reactions: *Journal of Experimental Psychology, 18,* 643-661.

Sussman, E. S., Bregman, A. S., & Lee, W. W. (2014). Effects of task switching on neural representations of ambiguous sound input. *Neuropsychologia, 64,* 218-229. doi: 10.1016/j.neuropsychologia.2014.09.039

Tamminen, J., Rastle, K., Darby, J., Lucas, R., & Williamson, V. (2015). The impact of music on learning and consolidation of novel words. *Memory, 29*(12), 1-1. doi: 10/1080/09658211.2015.1130843

Tanaka, S. M., & Kirino, E. (2017). Dynamic Reconfiguration of the Supplementary Motor Area network during Imagined Music performance. *Frontiers in Human Neuroscience*, *11,* 606. doi: 10.3389/fnhum.2017.00606

Thaut, M. H. (2009). The Musical Brain. An artful biological necessity. *Karger Gazette (Music and Medicine), 70,* 2-4.

Thomas, K. E. (1930). *The Real Personages of Mother Goose.* Lothrop, Lee and Shepherd Co. MW Books. ASIN: B0006AL375

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*(3), 248-251. http://dx.doi.org/10.1111/1467-9280.00345

Thompson, W. F., Schellenberg, E. G., & Letnic, A. K. (2011). Fast and loud background music disrupts reading comprehension. *Psychology of Music*, *40*(6), 700-708. doi: 10.1177/0305735611400173

Thompson, L. M., & Yankeelov, M. J. (2012). Music and the phonological loop. Article presented at The European Society for the Cognitive Sciences of Music, July 23-28, 2012. Thessaloniki, Greece.

Threadgold, E., Marsh, J. E., McLatchie, N., & Ball, L. J. (2019). Background music stints creativity: Evidence from compound remote associate tasks. *Applied Cognitive Psychology, 33,* 873-888. https://doi.org/10.1002/acp.3532

Tian, X., & Poeppel, D. (2012). Mental imagery of speech: linking motor and perceptual systems through internal similarities and estimation. *Frontiers in Human Neuroscience, 28*(6), 314. doi: 10.3389/fnhum.2012.00314

Tierney, A. T., Bergeson-Dana, T. R., & Pisoni, D. B. (2008). Effects of early musical experience on auditory sequence memory. *Empirical Musicology Review*, *3*(4), 178-186. doi: 10.18061/1811/35989

Tierney, A. T., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience, 8*(949), 1-9. doi: 10.3389/fnhum.2014.00949

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: a selforganizing approach. *Psychology Review 107*(4), 885-913. http://dx.doi.org/10.1037/0033-295X-107.4.885

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2008). Tonal centers and expectancy: Facilitation or inhibition of chords at the top of the harmonic hierarchy? *Journal of Experimental Psychology. Human Perception and Performance*, *34*(4), 1031-1043. doi: 10.1037/0096-1523.34.4.1031 2008

Tipper, S. P. (2001). Does Negative Priming Reflect Inhibitory Mechanisms? A Review and Integration of Conflicting Views. *The Quarterly Journal of Experimental Psychology*, *54*(2) 321-43. https://doi.org/10.1080/713755969

Tipper, S. P., & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *37*(4), 591-611. http://dx.doi.org/10.1080/14640748508400921

Tomasello, M., Velichkovskiĭ, B. M., & Rumbaugh, D. M. (1996). Universität Bielefeld. Zentrum für Interdisziplinäre Forschung, (Eds.), *The cultural roots of language. Communicating meaning: the evolution and development of language.* Mahwah, N. J.: L. Erlbaum. ISBN 978-0-8058-2118-5. OCLC 34078362.

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 659. doi: 10.1037/0278-7393.24.3.659

Tremblay, S., & Jones, D. M. (1999). Change of intensity fails to produce an irrelevant sound effect: Implications for the representation of unattended sound. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1005-1015. doi: 10.1037/0096-1523.25.4.1005

Tremblay, S., Macken, W. J., & Jones, D. M. (2001). The impact of broadband noise on serial memory: Changes in band-pass frequency increase disruption. *Memory, 9*(4-6*)*, 323-331. doi: 10.1080/09658210143000010

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, and Cognition, 26*(6), 1750-1754. doi: 10.1037/0278-7393.26.6.1750

Trollinger, V. L. (2010). The brain in singing and language. *General Music Today, 23*(2), 20-23. doi: 10.1177/1048371309353878

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*(6), 499-504.

Tsang, C. D., Friendly, R. H., & Trainor, L. J. (2011). Singing development as a sensorimotor interaction problem. *Psychomusicology: Music, Mind, and Brain, 21*(1-2), 31-44. http://dx.doi.org/10.1037/h0094002

Tucker, A., Bushman, B. (1991). Effects of rock and roll music on mathematical, verbal, and reading comprehension performance. *Perceptual and Motor Skills, 72*(3), 942. https://doi.org10.2466/pms.1991.72.3.942

Tulving, E., Kapur, S., Craik, F. I., Moscovitch, M., Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proceedings National Academy of Sciences USA, 91*(6), 1989-9. doi: 10.1073/pnas.91.6.2016

Tun, P. A., O'Kane, G., & Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging*, *17*(3), 453-467. http://dx.doi.org/10.1037/0882-7974.17.3.453

Ulbæk, I. B. (1998). The origin of language and cognition. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (Eds.), *Approaches to the evolution of language: social and cognitive base* (pp. 30-43). Cambridge, UK; New York: Cambridge University Press, ISBN 978-0-521-63964-4. OCLC 37742390.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by e procedural system. *Journal of Cognitive Neuroscience*, *9*(2), 266-276. doi: 10.1162/jocn.1997.9.2.266

Vachon, F., Hughes, R. W., & Jones, D. M. (2012). Broken expectations: Violation of expectancies, not novelty, captures auditory attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 164-177. doi: 10.1037/a0025054

Vachon, F., Labonté, K., & Marsh, J. E. (2017). Attentional capture by deviant sounds: A non-contingent form of auditory distraction? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 622-634. doi: 10.1037/xml0000330

Vachon, F., Marsh, J. E., & Labonté, K. (2020). The automaticity of semantic processing revisited: Auditory distraction by a categorical deviation. *Journal of Experimental Psychology: General,* Nov 21. doi: 10.1037/xge0000714 [Epub ahead of print.]

Van de Cavey, J., Severens, E., & Hartsuiker, R. J. (2017). Shared structuring resources across domains: Double task effects from linguistic processing on the structural integration of pitch sequence. *The Quarterly Journal of Experimental Psychology, 70*(8), 1633-1645. doi: 10.1080/17470218.2016.1195852

Van Den Bosch, I., Salimpoor, V., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience, 7*(534), 1-10. doi: 10.3389/fnhum.2013.00534

Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1471-1485. http://dx.doi.org/10.1037/0278-7393.20.6.1471

Wallin, N., Merker, B., & Brown, S. (2000). *The Origins of Music.* The MIT Press. ISBN: 0262 23206 5

Warren, R. (1961). Illusory changes of distinct speech upon repetition – the transformational effect. *British Journal of Psychology, 52,* 249-258. doi: 10.1111/j.2044-8295

Warren, R. (1999). *Auditory Perception: A new analysis and synthesis.* New York: Cambridge University Press.

Warren, R. M., & Gregory, R. L. (1958). An auditory analogue of the visual reversible figure. *American Journal of Psychology, 71*(3), 612-613.

Warren, R. M., & Obuzek, C. J. (1972). Identification of temporal order within auditory sequences. *Perception & Psychophysics, 12,* 86-90.

Watkins, M. J., & Allender, L. E. (1987). Inhibiting word generation with word presentations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*(4), 564-568. http://dx.doi.org/10.1037/0278-7393.13.4.564

Watson, A. H. D. (2006). What studying musicians tell us about motor control of the hand. *Journal of Anatomy, 208*(4), 527-542. doi: 10.1111/j.1469-7580.2006.00545.x

Weinstein, J., Koenig, P., Gunawardena, D., McMillan, C., Bonner, M., & Grossman, M. (2011). Preserved musical semantic memory in semantic dementia. *Archives of Neurology, 68*(2), 248-250. doi: 10.1001/archneurol.2010.364

Weiss, M. W., Trehub, S. E., & Schellenberg, E. G. (2012). Something in the way she sings: enhanced memory for vocal melodies. *Psychological Science, 23*(10), 1074-1078. doi: 10.1177/0956797612442552

Weiss, M. W., Trehub, S. E., Schellenberg, E. G., & Habashi, P. (2016). Pupils dilate for vocal or familiar music. *Journal of Experimental Psychology: Human Perception and Performance, 42*(8), 1061-1065. doi: 10.1037/xhp0000226

Whorf, B. L. (1956). *Language, Thought, and Reality. Selected writings of Benjamin Lee Whorf,* MIT Press, Paperback Series. ISBN-10:0262730065

Williamson, V. J., Baddeley, A. D., & Hitch, G. (2010a). Musicians' and non-musician's short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition, 38*(2), 163-175. doi:10.3758/MC.38.2.163

Williamson, V. J., Jilka, S. R., Fry, J., Funkel, S., Müllensiefen, D., & Stewart, L. (2011). How do earworms start? Classifying the everyday circumstances of involuntary musical imagery. *Sage Journals*, *40*(3), 259-284. doi: 10.1177/0305735611418553

Williamson, V. J., Mitchell, T., Hitch, G., & Baddeley, A. D. (2010b). Musicians' memory for verbal and tonal material under conditions of irrelevant sound. *Psychology of Music*, *38*(3), 331-350. doi: 10.1177/0305735609351918

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(14), 625-636. https://doi.org/10.3758/BF03196322

Wise, K. J., & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined "tone deafness": Perception, singing performance and self-assessment. *Musicae Scientiae, 12,* 3-26. http://dx.doi.org/10.1177/102986490801200102

Witt, J. K. (2011). Action's effect on perception. *Sage Publications, 20*(3), 201-206. https://doi.org/10.1177/0963721411408770

Wolfe, D. E., & Hom, C. (1993). Use of melodies as structural prompts for learning and retention of sequential verbal information by preschool students. *Journal of Music Therapy, 30*(2), 100-118. doi: 10.1093/jmt/30.2.100

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology*: *Learning, Memory, and Cognition, 21*(1), 255-60. http://www.ncbi.nlm.nih.gov/pubmed/7876773

Woodward, A. J., Macken, W. J., & Jones, D. M. (2008). Linguistic familiarity in short-term memory: A role for (co-)articulatory fluency? *Journal of Memory and Language, 58*(2008), 48-65. doi: 10.1016/j.jml.2007.07.002

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia, 23*(1), 31-41.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience, 8*(7), 547-58. doi: 10.1038/nrn2152

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14*(4), 1908-1919. doi: 10.1523/JNEUROSCI.14-04-01908.1994

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, *47*(1), 9-12. doi: 10.1016/jneuron.2005.06.013

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioural and neural correlates of executive functioning in musicians and non-musicians. *PLoS One, 9*(6), e99868. doi: 10.1371/journal.pone.0099868

Appendix A

Ethical approval for research project

20th January 2015

Dr. John E. Marsh/Rona D Linklater

School of Psychology

University of Central Lancashire

Dear John & Rona,

Re: PSYSOC Ethics Committee Application Unique Reference Number: PSYSOC 180

The PSYSOCS ethics committee has granted approval of your proposal application '**Music messaging for the brain: A study in Distraction, Retrieval and Recognition of Music**'. Approval is granted up to the end of project date* or for 5 years from the date of this letter, whichever is the longer. It is your responsibility to ensure that

- the project is carried out in line with the information provided in the forms you have submitted
- you regularly re-consider the ethical issues that may be raised in generating and analysing your data
- any proposed amendments/changes to the project are raised with, and approved, by Committee
- you notify roffice@uclan.ac.uk if the end date changes or the project does not start
- serious adverse events that occur from the project are reported to Committee
- a closure report is submitted to complete the ethics governance procedures (Existing paperwork can be used for this purposes e.g. funder's end of grant report; abstract for student award or NRES final report. If none of these are available use e-Ethics Closure Report Proforma).

Additionally, PSYSOCS ethics committee has listed the following recommendation(s) which it would prefer to be addressed. Please note, however, that the above decision will not be affected should you decide not to address any of these recommendation(s).

Should you decide to make any of these recommended amendments, please forward the amended documentation to roffice@uclan.ac.uk for its records and indicate, by completing the attached grid, which recommendations you have adopted. Please do not resubmit any documentation which you have **not** amended.

Yours sincerely,

Gayle Brewer Vice-Chair

PSYSOC Ethics Committee

* for research degree students this will be the final lapse date

Appendix B

Song questionnaire: Results

Please rate the song titles below as:

1 Very Familiar 2 Familiar 3 Vaguely Familiar 4 Unfamiliar

Participant Response x 20 as %	**%**	**%**	**%**	**%**	**3 + 4 < 55%**
Song Title x 80	**1**	**2**	**3**	**4**	**x**
Baa, Baa, Black Sheep	100	0	0	0	
Hot cross buns!	80	15	5	0	
London Bridge	85	15	0	0	
Here-we-go-round the mulberry bush	75	20	5	0	
Old King Cole	20	60	10	10	
Happy Birthday	100	0	0	0	
It’s raining, it’s pouring	5	10	20	65	x
Rock-a-bye Baby	65	25	10	0	
Three blind mice	100	0	0	0	
Georgie Porgie	5	0	10	75	x
Oranges and lemons	70	20	10	0	
A-tisket a-tasket	0	5	15	80	x
Pussy cat, pussy cat	30	10	25	35	x
Lavender’s Blue	60	25	10	5	
Mary, Mary, quite contrary	55	15	25	5	
I had a little nut tree	15	35	5	45	
Yankee Doodle	70	15	15	0	
God save the Queen	100	0	0	0	
Row, row, row your boat	20	35	25	20	
Ten green bottles	80	10	5	0	
Diddle, diddle, dumpling, my son John	0	0	10	90	x
Twinkle, twinkle, little star	100	0	0	0	
Jack and Jill	25	65	5	10	
The Miller of Dee	15	30	10	45	x
Auld Lang Syne	95	5	0	0	
Pop goes the weasel	85	5	5	5	
Ride a cock horse to Banbury Cross	5	10	25	60	x
Humpty Dumpty	95	0	5	0	
Rub-a-dub dub	0	5	10	85	x
Oh where, oh where has my little dog gone	35	15	15	35	
As going to St Ives	0	0	10	90	x
Jingle bells	95	5	0	0	
Lucy Locket	0	0	10	90	x
The wheels on the bus	85	10	5	0	
Dance to your daddy	75	5	0	20	
I love little pussy	20	5	0	75	x
Hey diddle diddle	45	5	30	20	
The Muffin man	10	10	25	55	x
Jack Sprat	0	5	20	75	x
Oh dear! What can the matter be?	70	25	5	0	
Sing a song of sixpence	80	10	5	5	
Little Jack Horner	15	25	35	25	x
Goosey, goosey gander	10	45	15	30	
Bobby Shafto’s gone to sea	15	35	10	40	

Ding-dong bell	0	0	10	90	x
The Farmer’s in the dell	40	30	10	20	
The sun has got his hat on	10	45	35	10	
Merrily we roll along	30	10	45	15	x
One, two, buckle my shoe	10	15	5	70	x
The grand old Duke of York	45	20	5	30	
Little Bo-Peep	15	40	20	25	
Little Robin Redbreast	0	0	5	95	x
One for sorrow	0	5	25	70	x
Pat-a-cake	10	55	15	20	
Doctor Foster	5	0	10	85	x
Ring-a-ring o’ roses	55	20	15	10	
Old MacDonald had a farm	85	5	10	0	
Simple Simon	30	5	10	55	x
One man went to mow	30	15	30	25	x
This old man	60	15	15	10	
There was an old woman who lived in a shoe	0	15	30	55	x
Pease pudding hot	10	0	25	65	x
Hush little baby	0	0	5	95	x
Little Miss Muffett	10	65	5	20	
There was a crooked man	0	15	45	40	x
For he’s a jolly good fellow	65	15	20	0	
Wee Willy Winky	0	25	40	35	x
If your happy and you know it	70	15	15	0	
Little Tommy Tucker	0	0	10	90	x
I had a little teapot	5	0	0	95	x
Itsy bitsy spider	15	30	20	35	x
Hickory, dickory, dock	10	45	15	30	
Needles and pins	0	0	25	75	x
Old Mother Hubbard	5	25	30	40	x
What are little boys made of	0	0	35	65	x
Girls and boys come out to play	10	45	30	15	
Lucy Locket	0	15	20	65	x
Little boy blue	10	45	25	20	
Polly, put the kettle on	25	35	10	30	
Oh, when the saints	30	55	10	5	

Total > 55%

47 Titles

Thank You

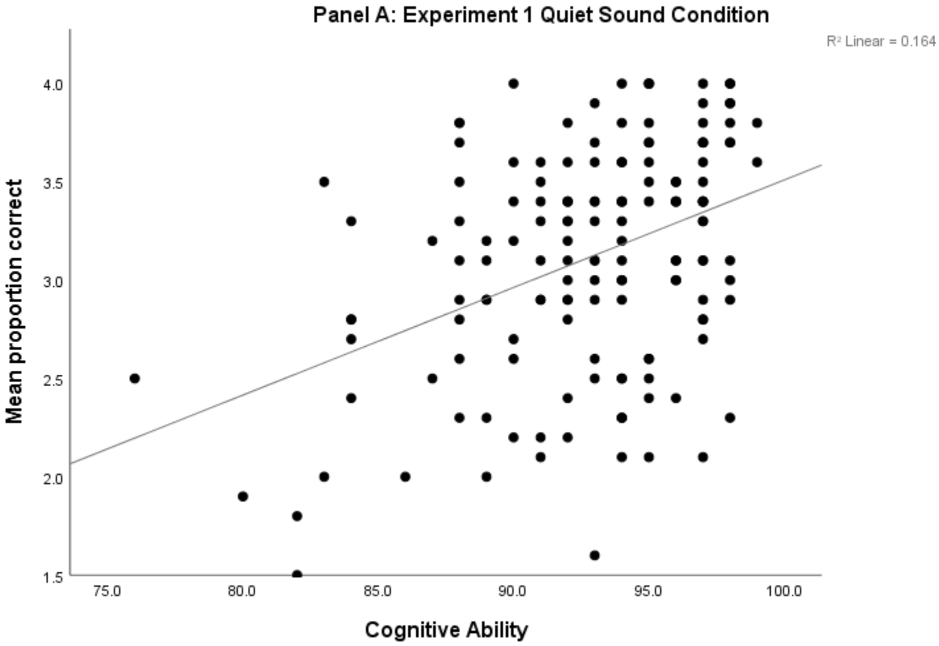
Appendix C

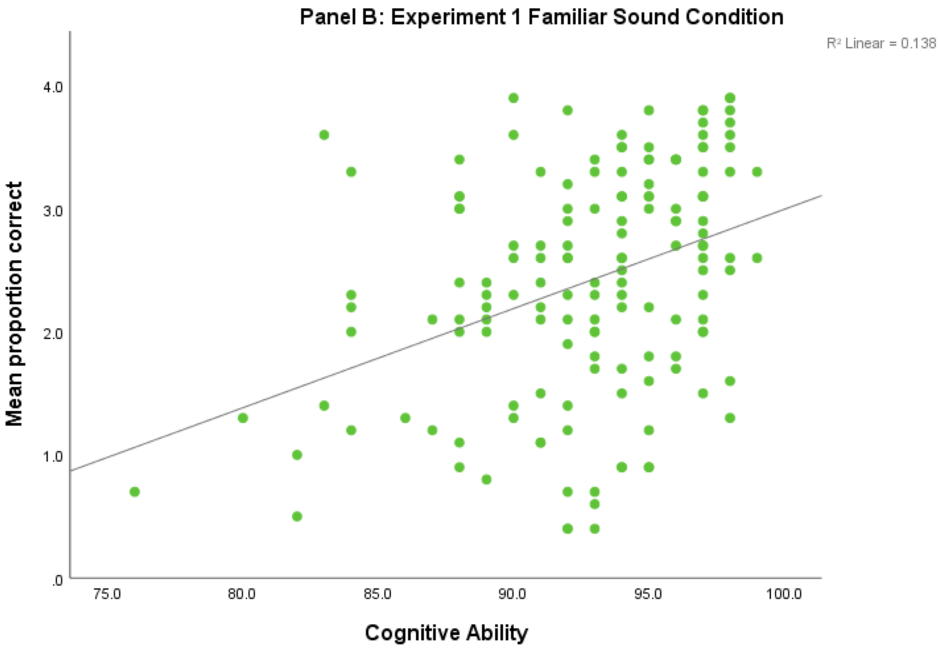
Song Sets

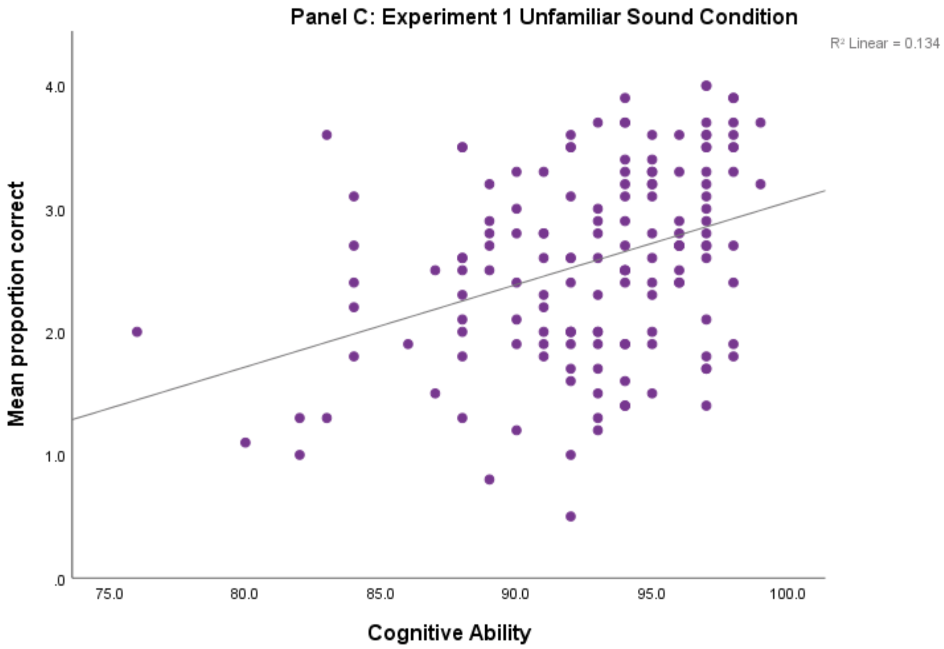
A	sec	B	sec	C	sec	D Distracter	sec
Baa, Baa, black sheep	16	Hot cross buns!	16	Old King Cole	16	London Bridge	16
The Mulberry Bush	24	Hickory, dickory, dock	24	Rock-a-bye Baby	24	Pop goes the weasel	24
Little Bo-Peep	24	Pat-a-cake	24	Girls and boys come out to play	24	Little Boy Blue	24
Polly, put the kettle on	16	Old MacDonald had a farm	16	This old man	16	Little Miss Muffet	16
Sing a song of sixpence	16	The grand old Duke of York	16	Mary, Mary quite contrary	16	If you're happy	16
Three blind mice	16	Auld Lang Syne	16	Goosey, Goosey, Gander	16	Oh, when the saints	16
Oranges and lemons	12	Lavender's Blue	12	Ten green bottles	16	Dance to your Daddy	12
Oh dear! What can the matter be?	24	Hey diddle, diddle	24	Jack and Jill	24	Oh where, oh where has my little dog gone?	24
God save the Queen	21	Twinkle, twinkle, little star	24	Jingle Bells	16	Jolly good fellow	20
The wheels on the bus	16	Happy Birthday	12	The Farmer's in the dell	16	Yankee Doodle	16
Total Humming / Total Speaking Seconds	**185**		**184**		**184**		**184**

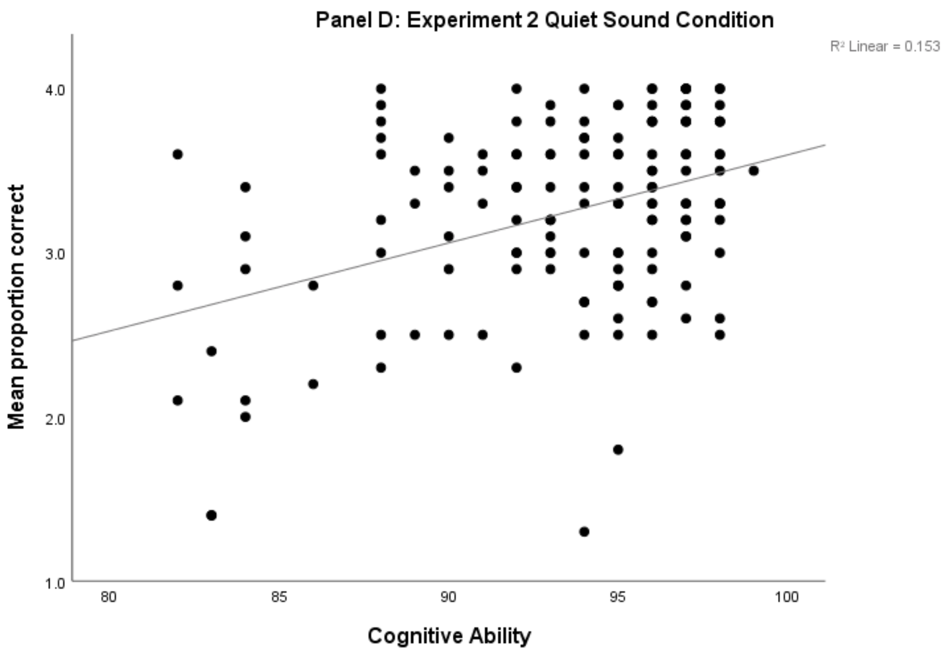
Appendix D

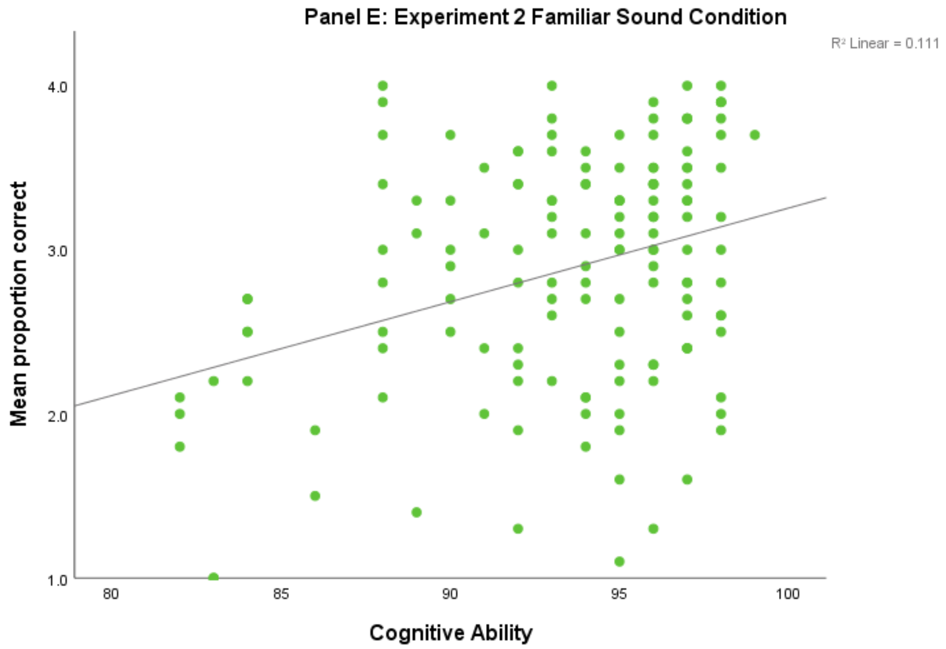
Recall scores *versus* cognition for Experiment 1, 2, and 3.

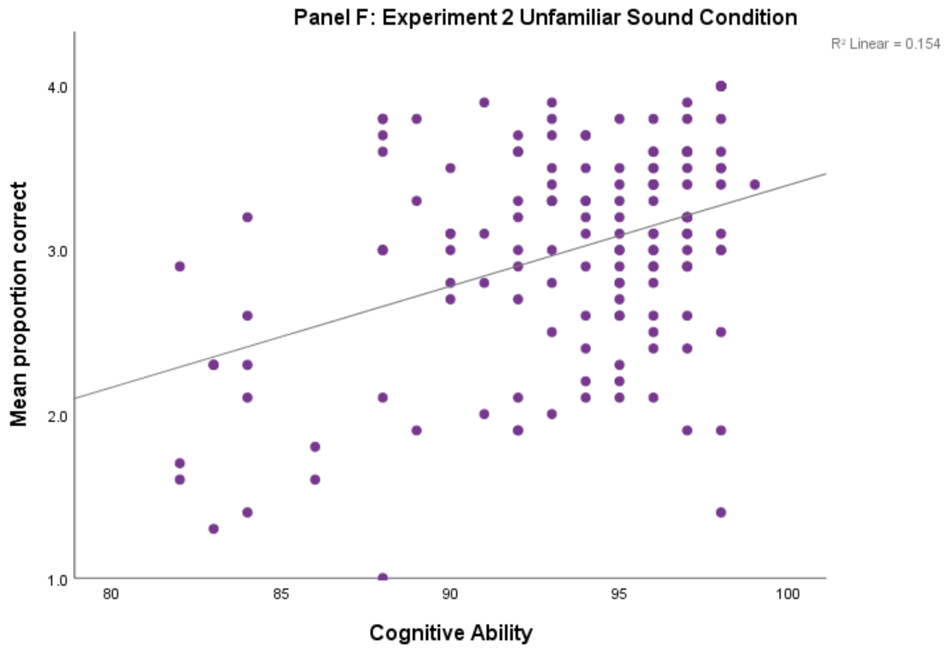


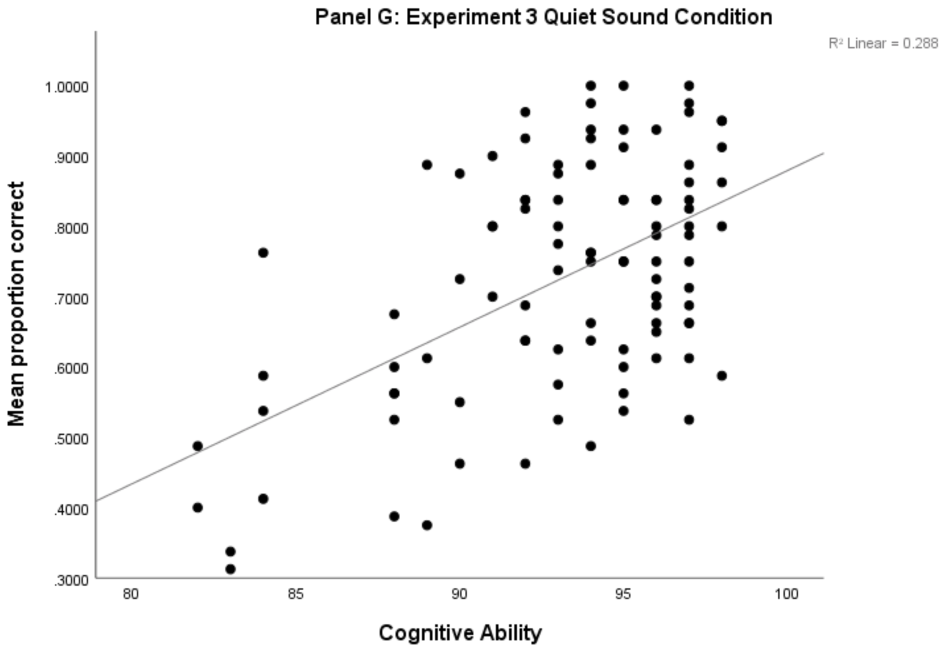


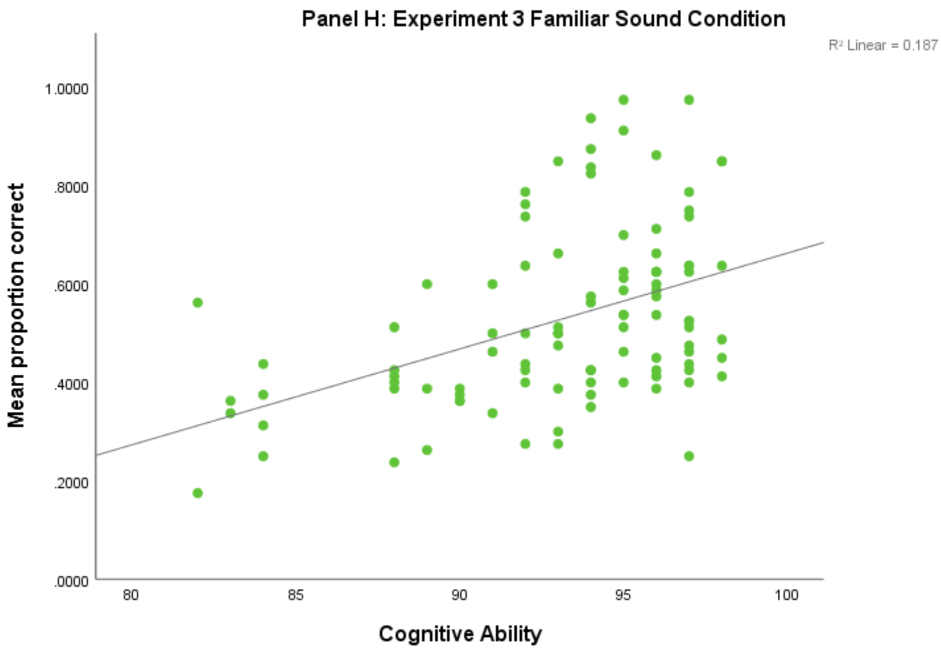

Panel C: Experiment 1 Unfamiliar Sound Condition
R² Linear = 0.134
Mean proportion correct
4.0
3.0
2.0
1.0
.0
75.0
80.0
85.0
90.0
95.0
100.0
Cognitive Ability

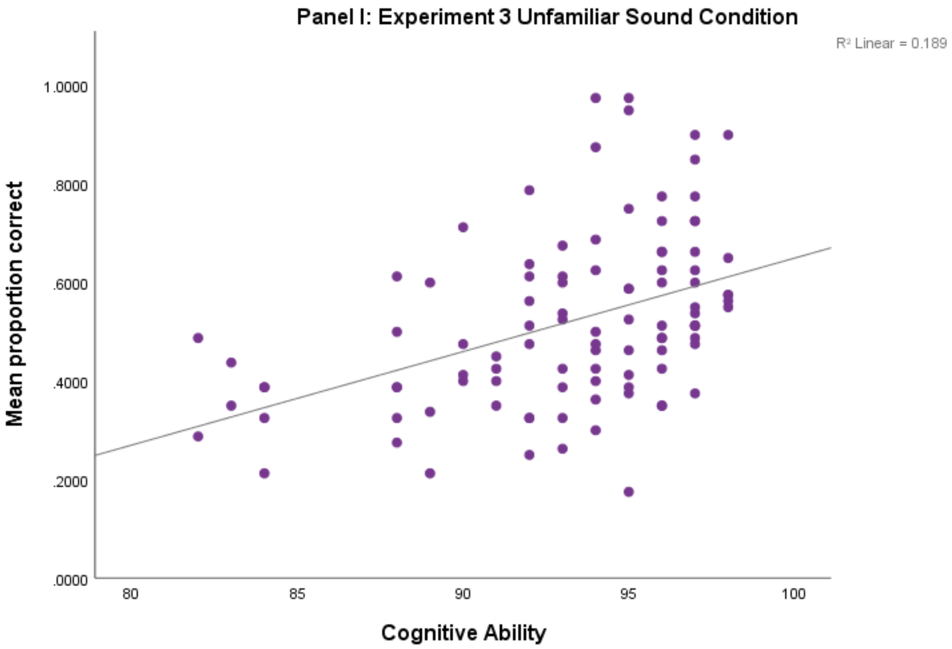

Panel D: Experiment 2 Quiet Sound Condition
R² Linear = 0.153
Mean proportion correct
4.0
3.0
2.0
1.0
80
85
90
95
100
Cognitive Ability


Panel E: Experiment 2 Familiar Sound Condition
R² Linear = 0.111
Mean proportion correct
4.0
3.0
2.0
1.0
80
85
90
95
100
Cognitive Ability


Panel F: Experiment 2 Unfamiliar Sound Condition
R² Linear = 0.154
Mean proportion correct
4.0
3.0
2.0
1.0
80
85
90
95
100
Cognitive Ability


Panel G: Experiment 3 Quiet Sound Condition
R² Linear = 0.288
Mean proportion correct
1.0000
.9000
.8000
.7000
.6000
.5000
.4000
.3000
80
85
90
95
100
Cognitive Ability


Panel H: Experiment 3 Familiar Sound Condition
R² Linear = 0.187
Mean proportion correct
1.0000
.8000
.6000
.4000
.2000
.0000
80
85
90
95
100
Cognitive Ability


Panel I: Experiment 3 Unfamiliar Sound Condition
R² Linear = 0.189
Mean proportion correct
1.0000
.8000
.6000
.4000
.2000
.0000
80
85
90
95
100
Cognitive Ability

Appendix E

Missing-item task: Recall strategy statements

Please indicate the strategy that you used to complete the digit recall task - press the key that corresponds to the strategy:

A	I thought about other things that could relate to the digits.
B	I answered based on what digits seemed recent or familiar.
C	I expected certain digits to appear and mentally checked them off as they arrived.
D	I silently repeated the items.
E	I remembered the items in groups.
F	I used the meaning of digits to connect them.
G	I pictured the way the digits looked on the screen.
H	I thought about how the digits sounded.
I	I simply concentrated on the digits.
J	I used an unidentified strategy.
K	I didn't understand the task instructions.