Title: Investigating the phonological similarity effect: implications for short-term memory models

Author: Lisa M Nimmo

Year: 2004

Repository DOI:

University of Wollongong *Year* 2004

Investigating the phonological similarity effect: implications for short-term memory models

Lisa M. Nimmo
University of Wollongong

INVESTIGATING THE PHONOLOGICAL SIMILARITY EFFECT: IMPLICATIONS FOR SHORT-TERM MEMORY MODELS

A thesis submitted in fulfilment of the requirements for the award of the degree

DOCTOR OF PHILOSOPHY

From

UNIVERSITY OF WOLLONGONG

By

Lisa M. Nimmo

BPsyc(Hons)

Department of Psychology

2004

CERTIFICATION

I, Lisa M. Nimmo, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Department of Psychology, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

Lisa M. Nimmo
9th January 2004

Acknowledgements

I would like to thank the members of the Psychology department at the University of Wollongong, who have helped me over the last three years. I love this department because of the friendships that I have formed. From the general staff (Lisa, Dayna & Cathy) where nothing is too difficult or a problem, to the academic staff (Amy Chan, Stephen Palmisano and Doug Cornford) that have offered me words of encouragement, praise, wisdom, or thrown a party just because everyone was feeling a little low. I thank you. In particular, I would like to thank Simone Keane who was always one step ahead. Whether it was tips on presentation style, the best way to format a thesis, or agreeing that 2:00 on any cricket day is a good time to go to the bar for a few beers. You have made this journey a lot easier for me.

I am grateful to my supervisor Steven Roodenrys. You are inspirational. You have instilled in me traditional research values. For you, and now myself, research is all about quality rather than quantity. I came to you when I was still an undergraduate student and said "cognitive psychology is for me". Since then you have taken me under your wing, let me explore my own research ideas and have guided me through the Honours, PhD and publication process. Although I have jumped in the deep end on a number of occasions, you always made sure that my floaties were securely fastened. Most importantly, I thank you for always having faith in my ability, which in turn made me believe.

Thanks also to my parents, Ray and Jean, for their constant love and support and for caring about all that "Uni stuff", just because I care about it.

Finally, I wish to thank Daniel Hughes. We have been in this together from start to finish. I thank you for your endless love and support and for your unwavering belief in my ability to succeed, not only in academia but also with anything I put my mind to. I couldn't have done it without you.

Abstract

The current thesis examined the effect that phonological similarity has on short-term memory (STM) performance. Across nine experiments, the predictions that two classes of STM models (non-linguistic and psycholinguistic) generate for the effect that phonological similarity has on the recall of item information and memory for an item's position in a list were tested.

In the current thesis, phonological similarity was operationally defined in a number of different ways. For instance, lists of consonant-vowel-consonant (CVC) words and nonwords, rhymed (shared _VC component), shared the initial consonant and vowel (CV_ component) or shared the two consonants (C_C component). Performance across these conditions was compared to when the stimulus lists were either phonemically dissimilar (i.e., used as a baseline measure of performance) or phonemically similar (i.e., each stimulus in each list had at least two phonemes in common with at least one other stimulus in the same list).

Regardless of whether the experimental stimuli were words or nonwords, when performance was measured using the item recall criterion (scored as correct if a participant recalled an item that was presented in a list, regardless of position), an item recall advantage was observed for rhyming lists of stimuli. Non-linguistic STM models suggest that an item recall advantage should be observed whenever the size of the 'secondary memory search set' can be limited to a smaller number of items (e.g., all items that rhyme). In contrast, psycholinguistic models of STM assume that this item recall advantage derives from sub-syllabic structures that aid the recall of item information.

In terms of the effect that phonemic similarity has on order memory, the findings from the current thesis are inconsistent with the predictions generated from non-linguistic models of STM that are based on the distinctiveness assumption – the idea that as similarity increases order memory should decrease. Rather the findings are consistent with psycholinguistic models of STM that assume that the effect that phonemic similarity has on order memory is a consequence of linguistic constraints, such as sonority, that operate at the sub-syllabic as compared to lexical level. Based on the current research findings, modifications to existing STM models have been proposed.

Table of Contents

List of Tables

List of Figures

Preface

Chapter 4 of this thesis has been published as part of a jointly authored paper:

> Nimmo, L.M., & Roodenrys, S. (in press). Investigating the phonological similarity effect: The Syllable structure and the position of common phonemes. *Journal of Memory and Language.*

Chapter 5 of this thesis is currently under review as part of a jointly authored paper:

> Nimmo, L.M., & Roodenrys, S. (submitted). Investigating the phonological similarity effect: Are nonwords an alien form? *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Chapter 6 of this thesis is currently under review as part of a jointly authored paper:

> Nimmo, L.M., & Roodenrys, S. (submitted). The phonological similarity effect in serial recognition. *Memory*.

For all three manuscripts, I designed the experiments, created the stimulus lists, programmed the experiments, recruited participants, collected and analysed the data, and prepared the manuscripts.

Parts of the research in this thesis were presented at: *The 30th Annual Conference of the Australasian Experimental Psychology Society,* University of Western Sydney, Sydney, NSW, Australia, April, 2003. *The Quebec'02 Conference on Short-term/Working Memory,* Quebec City, Canada, July, 2002. *The 6th Conference of the Australasian Cognitive Science Society,* Freemantle, Western Australia, April, 2002.

Synopsis

The effect that phonological similarity has on our ability to recall items from short-term memory (STM) is one of the theoretically most influential findings in studies of STM: This is the finding that serial recall performance is worse if words sound similar to each other (e.g., Conrad & Hull, 1964). However, when performance was measured for item recall (i.e., number of items recalled, regardless of position), Wickelgren (1965d) found no differences between phonemically dissimilar and similar lists of items. This lead earlier researcher's to conclude that phonological similarity influences the order in which items are recalled rather than the retention of item information (Murdock, 1976).

The effect that phonological similarity has on a participant's ability to recall list items in the correct order is such a robust finding in the STM literature that some researchers have suggested that the value or worth of extant STM models can be gauged by the explanations they generate for this effect (Gathercole, 1997; Nairne, 1990a; Page & Norris, 1998). As Nairne and Kelly (1999; p.45) suggest,

> *"…the phonological similarity effect has achieved the status of a 'benchmark' finding in the immediate memory literature, and most theories of short-term memory include mechanisms that are specifically designed to account for the phenomenon"*

However, recent research findings have questioned the stability of the phonological similarity effect. Although the detrimental effect that phonological similarity has on order memory has been replicated in numerous studies (e.g., Baddeley, 1966), when the effect that phonological similarity has on the recall of item information is examined, the results are contradictory. For instance, although some studies have found no differences between phonemically similar as compared to dissimilar lists of items (e.g., Poirier & Saint-Aubin, 1996), others have found that phonemic similarity can either facilitate (e.g., Fallon, Groves & Tehan, 1999) or have a detrimental effect (e.g., Coltheart, 1993) on the recall of item information.

A number of suggestions have been proposed to account for the contradictory findings that have recently been observed in the research literature. For instance, according to Fallon et al., (1999) differential results are observed in the literature depending on how phonological similarity has been operationally defined, the size of the word pools used to construct the stimulus lists, and the scoring criteria (i.e., correct-in-position, item recall, or order accuracy) used to measure STM performance.

In light of the inconsistencies that have recently been reported in the research literature, a major aim of the current thesis was to examine the effect that operationally defining similarity in different ways has on the recall of item information and memory for an item's position in a list. This was achieved by constructing lists of consonant-vowel-consonant (CVC) items that either shared the rhyme (_VC), the initial consonant and vowel (CV_), or the two consonants (C_C). Thus, the position of the overlapping phonemes was manipulated, while the amount of phonemic overlap (as measured by the degree of shared consonant and vowel information) was held constant. Performance on these types of lists was compared to when the stimulus lists were composed of either phonemically similar (i.e., each stimulus in each list consisted of at least two phonemes in common with at least one stimulus in the same list) or phonemically dissimilar (i.e., no item in a list shared any common phonemes with any other item in the same list) items.

A further aim of the current thesis hinged on the idea that "…any plausible model of short-term memory must explain" the phonological similarity effect (Lian, Karlsen & Winsvold, 2001; p.281). Currently, there are two distinct classes of STM models that attempt to provide an explanation for the effect that phonological similarity has on STM performance: psycholinguistic and non-linguistic models of STM. Psycholinguistic models of STM (e.g., Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) are based on the idea that the effect that phonemic similarity has on item and order memory derives from the influence that sub-syllabic linguistic mechanisms, such as syllable structure and sonority, have on STM performance. In contrast, non-linguistic STM models (e.g., Brown, Preece & Hulme, 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a, 2002) are based on the distinctiveness assumption – the idea that as similarity increases order memory should decrease. Thus, according to these types of models, if phonological similarity is held constant across experiments, similar levels of order memory impairment should be observed. Hence, the current thesis was designed to critically evaluate the utility of psycholinguistic and non-linguistic STM models by the explanations they generated for the effect that operationally defining similarity in different ways has on the recall of item information and memory for an item's position in a list.

The current thesis can be divided into three distinct sections, the first of which is three introductory chapters. Chapter one was designed to provide a broad overview of STM, how STM has traditionally been measured, and more general research findings related to the effect that both phonological similarity and lexicality have on STM performance. Chapter two was dedicated to describing the assumptions that STM models are based on, and more generally, the mechanisms that researchers incorporate into these models to account for a variety of STM research findings. The final introductory chapter (Chapter 3) critically examined the existing research into the effect that phonological similarity has on STM performance with a particular emphasis on the inconsistencies that have been found in the research literature and its relation to both the lexicality of the experimental items and the effect that overt speech production has on STM performance.

The second section of the current thesis consists of three experimental chapters. Each experimental chapter has been written in manuscript format[1] and are self-contained, in that they were designed to investigate different issues with respect to the effect that similarity has on STM performance (although all of the experiments were designed to examine the utility of STM models by the explanations they generate for the effect that phonological similarity has on STM performance). The aim of study one (Chapter 4 - Experiments 1 to 3) was to examine the effect that operationally defining phonemic similarity in different ways has on the recall of item information and memory for an item's position in a list when the experimental stimuli were words. Study two (Chapter 5 - Experiments 1 to 3) was designed to further examine this issue, but with nonwords as compared to words. This type of investigation is warranted in that to date, a number of STM models do not provide an explanation for the effect that the phonemic similarity of nonwords has on STM performance. This stems from the belief that

---

[1] Please note that although the wording has not changed for the manuscripts that are either in press or under review, the format has been changed to make these manuscripts consistent with the format that has been used in the current thesis.

"…given that no adequate long-term representations are available for nonwords, the reconstruction process, for all practical purposes, is thought not to operate for these items" (Saint-Aubin & Poirier, 2000; p.333). Finally, Gathercole, Service, Hitch, Adams and Martin (1999) have recently suggested that the findings observed from studies that require participants to verbally recall presented list items, may be influenced by an individual's articulatory ability, especially when the experimental stimuli are nonwords. Hence, regardless of whether the experimental stimuli were words or nonwords, study three (Chapter 6 - Experiments 1 to 3) was designed to examine the effect that phonemic similarity has on order memory, once the demands that overt speech production have on STM performance are removed.

The final section of the current thesis consists of two concluding chapters. Chapter seven draws a number of clear conclusions that are based on the current research findings. Firstly, the findings from the current thesis suggest that the same mechanisms are involved both word and nonword recall. Secondly, that the effect that similarity has on order memory remains, once the demands that overt speech production have on STM performance are removed. Finally, that STM models that are based on the distinctiveness assumption (e.g., Nairne, 1988, 1990a) are unable to account for the current research findings. Rather, the findings are more consistent with the explanations that psycholinguistic models of STM (e.g., Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) generate for the effect that phonological similarity has on the recall of item information and memory for an item's position in a list. The current thesis culminates (Chapter 8) with an in-depth discussion of the implications that the current research findings have for extant STM models with a particular emphasis on modifications to existing STM models and suggestions for future research.

1. SHORT-TERM MEMORY

1.1. Short-term memory: A brief overview

> *"Primary memory, elementary memory, immediate memory, short-term memory (STM), short-term store (STS), temporary memory, supervisory attention system (SAS), working memory (WM) – all these terms refer to the same memory component, the same aspect of the human information-processing system" (Ashcraft, 2002; p. 160)*

It has become somewhat of a tradition to define STM in relation to long-term memory (LTM). Whereas STM holds information temporarily and is limited in capacity (Miller, 1956), LTM is neither limited in capacity nor temporally constrained (Baddeley, 1997). Although information that is attended to, such as rehearsing a phone number that you are about to call is held in STM, information about what you ate for dinner the previous night is recalled from LTM (Galotti, 1999). While memory theorists drew a distinction between LTM and STM more than a century ago (e.g., William James, 1890), in the 1960s a great debate raged as to whether STM was separable from LTM (Atkinson, 1968; Melton, 1963). Nowadays it is a widely held belief that STM and LTM are two functionally distinct systems (Baddeley & Logie, 1999; Burgess & Hitch, 1999). However, this is not to suggest that long-term knowledge does not influence STM performance. Rather, current empirical evidence suggests that LTM does indeed influence performance on STM tasks (see Baddeley & Logie, 1999 for an in-depth discussion of the empirical evidence in support of this view).

1.2. How is short-term memory measured?

The first direct attempt at measuring STM was conducted by Joseph Jacobs (1887). The measurement of STM arose from Jacob's interest in assessing the mental capacity of his students (Baddeley, 1997). The technique that he devised to measure STM capacity has been termed the *memory span procedure* (Wechsler, 1991, 1997). In this task, participants are presented with a list of items and are asked to repeat them back in the order in which they were presented. The number of items in a presented list increases in a step-like fashion until a participant consistently fails to recall a list at a particular length (Jacobs, 1887). An individual's memory span is typically calculated at the point at which he/she correctly recalls 50% of the lists at a certain length (Cowan, 1992; Cowan et al., 1994; Cowan, Saults, Winterowd & Sherk, 1991; Cowan et al., 1998; Gathercole et al., 1999).

Since the memory span procedure was developed, a number of different techniques have been created to examine STM. The use of a particular STM task is constrained by the purpose of the test. In general, there are two ways in which STM tasks vary. The type of experimental stimuli (e.g., letters, words, nonwords, varying in length, phonological similarity or semantic similarity) can be manipulated to assess whether stimulus differences influence STM performance (e.g., Baddeley, Thomson & Buchanan, 1975; Conrad & Hull, 1964; Hulme, Maughan & Brown, 1991; Poirier & Saint-Aubin, 1995). Alternatively, researchers (Bjork & Whitten, 1974; Lian et al., 2001; Murdock, 1968; Poirier & Saint-Aubin, 1996) can vary the task demands (e.g., immediate serial recall task, free recall, probed recall or a serial recognition task) to investigate whether differences in the stimuli used across studies influences the recall of item information or memory for an item's position within a list.

1.2.1. The standard immediate serial recall (ISR) task

One of the most commonly used measures of verbal STM ability is the immediate serial recall (ISR) task (Gupta & MacWhinney, 1997). Immediately following the presentation of a sequence of items, a participant is required to recall the items in the order in which they were presented (Ashcraft, 2002). Due to the way in which performance is measured, this task can be used to assess both the recall of item information, and a participant's memory for an item's position in a list. For instance, performance can be measured using a *correct-in-position criterion* (i.e., scored as correct if a participant recalls the correct item in the correct position). This yields a measure of the number of items recalled in the correct order (Conrad & Hull, 1964). Performing well on such a task requires that participant's remember both item and order information. Performance can also be measured using an *item recall criterion* (i.e., scored as correct if a participant recalls an item presented in a given list, regardless of position). This scoring criterion yields a measure of the number of items recalled that is not dependent on memory for an item's position within a list (Wickelgren, 1965d). Hence, to perform well using this measure, a participant has to remember the items that were presented rather than the order in which they occurred (Wickelgren, 1965d).

Researchers have suggested that differences in item recall performance across different conditions influences the absolute number of order errors obtained (Murdock, 1976; Saint-Aubin & Poirier, 1999a, 1999b). For instance, no order errors are possible if a participant does not recall any items. Hence, when performance is scored using the item recall criterion, as the number of items recalled increases so do the number of possible order errors. To control for the influence that individual differences in item recall have on the number of order errors, researchers (Fallon et al., 1999; Saint-Aubin & Poirier, 1999a, 1999b) suggest that performance should be measured using a third scoring criterion - an *order accuracy measure*. The order accuracy measure is obtained by dividing the score observed using the correct-in-position criterion by the score obtained when performance is measured using the item recall criterion. This yields a measure of the proportion correct as a function of the number of items recalled. As such, the order accuracy criterion provides a measure of order memory that takes into account differences in item availability between conditions (Saint-Aubin & Poirier, 1999a, 1999b).

1.2.2. Item information

A task that is used to assess item information independent of order constraints, is called the *free recall task* (Bjork & Whitten, 1974; Glanzer & Cunitz, 1966; Postman & Phillips, 1965). In free recall, participants are presented with a list of items and their task is to recall as many of the items as possible in any order (Atkinson & Shiffrin, 1971; Craik, 1970). The dependent measure is typically the proportion of items recalled.

1.2.3. Order information

There are also a number of tasks that have been designed to assess a participant's memory for positional information (Healy, 1974; although Neath, 1997, argues against the idea that order and item information are separable; see section 3.10.1). For instance, an *order reconstruction task* involves presenting participants with a list of items one at a time and then presenting the same items simultaneously, either alphabetically or in a random order (Healy, Fendrich, Cunningham & Till, 1987; Nairne, 1991, 1992; Nairne & Neumann, 1993; Neath, 1997). A participant's task is to reassemble the previously presented list so that the items are placed in the order in which they were originally presented (Neath, 1997). The dependent measure when performing an order reconstruction task is usually the percentage of items placed in the correct position (Nairne, 1991).

A *probed recall task* is similar to the order reconstruction task except that after a list has been presented to a participant, one item that is called a cue or probe acts as a recall prompt (Waugh & Norman, 1965). A participant's task is to recall the item that succeeded the probe in the presented list (Murdock, 1968). The dependent measure on this type of task is commonly the mean proportion correct (Avons, Wright & Pammer, 1994). A third type of STM task that is designed to assess memory for positional information is the *serial recognition task* (Campbell & Butterworth, 1985; Martin & Breedin, 1992; Martin, Lesch & Bartha, 1999). This task involves presenting a list of items to participants and then re-presenting either the same item's in the same order or the same item's in a different order (Lian et al., 2001). Hence, this is a forced choice paradigm in which a participant's task is to say whether the items were presented in the 'same' or a 'different' order (Gathercole, Pickering, Hall & Peaker, 2001). As such, the dependent measure on this task is the proportion of trials correct (Gathercole et al., 2001).

1.2.4. Short-term memory measures: Weighing up the advantages and disadvantages

Researchers should be aware of both the benefits and limitations that arise when using different types of STM measures. For instance, as compared to STM performance when measured using the probed recall task, performance on the order reconstruction task is influenced to a greater extent by guessing (Healy, 1974). This is because participants are given all of the item information that was presented to them, *during* the recall phase of the experiment. Therefore, when reconstructing a five-item list, participants need only remember the order of the first four list items to correctly order the fifth item. However, one disadvantage of using a probed recall task to measure STM performance is that each list yields a score out of one. In contrast, when the task is order reconstruction each correctly ordered item in a list is scored (Nairne, 1991), thus providing a more sensitive measure of the particular STM effect under investigation.

There are also a number of advantages to using the serial recognition as compared to the ISR task. For instance, because the recognition task is not as demanding as the standard ISR task, this measure is less stressful. As such, the results obtained from recognition tasks are less prone to the effect that test anxiety may have on STM performance (Clegg & Warrington, 1994). Recently, Gathercole et al. (2001) have suggested that it may be a better measure of STM because the serial recognition task is resistant to speech production errors. For instance, unlike the standard ISR task, the serial recognition task does not require participants to overtly articulate the presented list items (Martin et al., 1999).

However, there are also a number of disadvantages to using a serial recognition task as compared to the standard ISR task. The first disadvantage is that when a list consists of five items, for the serial recall task each list is scored out of five. Conversely, when the task is serial recognition there is only one data point per list (i.e., a participant is either correct or incorrect). Hence, because serial recognition, like the serial probed task, is not as sensitive a measure of performance as the standard ISR task, a greater number of participants (or alternatively, a greater number of trials per participant) may be required to reach the same level of statistical power (Gathercole et al., 2001). Also, a richer body of data can be gathered using the standard ISR task as compared to any of the previously mentioned STM measures. For instance, the different types of errors that occur when performing a standard ISR task can be collated (Conrad, 1965; Ellis, 1980; Treiman & Danis, 1988). Finally, performing well on the standard ISR task requires a deeper level of information processing than is required to perform well on a recognition task (Ashcraft, 2002; Sternberg, 1999). In other words, it is unclear as to whether participants encode parts or all of the items presented when performing either a serial recognition or order reconstruction task.

Gathercole and McCarthy (1994) suggest that if STM tests are to be useful, a minimum of two criteria need to be satisfied. The first is reliability - do individuals that score well on a particular STM task consistently score well on that task? The second is concerned with validity - does a particular task measure what it purports to measure? (Murphy & Davidshofer, 1998) The methods used to assess STM have been replicated on numerous occasions with different variations. As such, it is fair to say that these measurement techniques themselves are no longer subject to questions about reliability and validity. However, what is subject to these questions is whether a particular researcher has chosen the most appropriate STM measure to answer the research question posed.

1.3. The phonological similarity effect (PSE)

A densely researched area in the STM literature is the effect that phonological similarity has on STM performance. In a pioneering study, Conrad and Hull (1964) examined whether the acoustic confusability of visually presented letters influenced written recall performance. They found that when the experimental stimuli sounded similar, performance was lower in comparison to when the lists were composed of distinct sounding letters. This early research finding has been replicated and extended to include other presentation techniques, recall methods and verbal materials (e.g., Baddeley, 1966; Coltheart, 1993; Cowan et al., 1991; Fallon et al., 1999; Gathercole, Gardiner & Gregg, 1982; Li, Schweickert & Gandour, 2000; Nairne, 1990a; Nairne & Kelley, 1999; Wickelgren, 1965b, 1965d). This robust STM finding has been termed the *phonological similarity effect* (PSE).

STM modellers have suggested that the worth of STM models lie in their ability to account for the PSE amongst other effects (Nairne & Kelley, 1999). As such, a multitude of STM models have attempted to provide an explanation for the PSE (e.g., Baddeley, 1986; Brown et al., 2000; Burgess & Hitch, 1992, 1999; Gupta & MacWhinney, 1997; Nairne, 1990a; Page & Norris, 1998; Schweickert, 1993). Although the PSE is considered to be a robust finding in the STM literature, as the number of studies investigating this effect increase, so do the smaller number of contradictory results (Coltheart, 1993; Gathercole et al., 1982; Watkins, Watkins & Crowder, 1974). Three suggestions have been proposed to account for these discrepant findings. For instance, Murdock (1976; see also Poirier & Saint-Aubin, 1996) suggested that the effect that similarity has on STM is dependent on the scoring criterion used to measure recall performance. In addition, Coltheart (1993) suggested that the differential results observed in the literature may depend on whether the stimulus lists were chosen using a closed (i.e., words that are sampled repeatedly from the same small set) or open word pool (i.e., words that are sampled without replacement from a large set). Finally, Fallon et al. (1999) have found contradictory results in the research literature when similarity is operationally defined in different ways. Hence, given the importance of research findings into the PSE for STM modelling, the current thesis has identified an urgent need to examine the contradictory findings reported in the literature on the effect that similarity has on STM performance.

1.4. A theoretical framework: The working memory model

Baddeley and Hitch's (1974) working memory model has been influential over the last few decades as a theory of STM that has guided research within the cognitive sciences. The working memory model consists of a central executive and two slave systems: the *visuo-spatial sketchpad* and the *phonological loop* (PL) (Baddeley & Hitch, 1974; Baddeley, 1986). The central executive is responsible for controlling attention and coordinating the activities of the other two components. The visuo-spatial sketchpad is responsible for both the maintenance and manipulation of spatial and visual images, whereas the PL is specialised in that it retains verbal information for short periods of time. The PL has a further two sub-components: a *phonological store* and an *articulatory rehearsal mechanism*. Whereas the store holds phonological information in the form in which it was heard, the rehearsal mechanism is responsible for maintaining decaying information in the store (Baddeley, 1986).

There are four experimental findings that are generally used to support the idea of a PL. The first is the effect that phonological similarity has on STM performance. As described previously, this is the finding that it is harder to recall lists of similar as compared to distinct sounding items (Baddeley, 1966; Conrad & Hull, 1964). This finding supports the idea that the phonological store is speech based (Andrade, 2001). A second experimental observation that is used as evidence for the PL is the word length effect. The word length effect is the finding that performance decreases as the length of the to-be-recalled item's increase (Baddeley et al., 1975; Schweickert & Boruff, 1986), hence lending support to the idea that STM is limited by an individual’s articulatory speed and thus rate at which items held in the phonological store can be refreshed through the use of the rehearsal mechanism (Burgess & Hitch, 1999). Thirdly, the irrelevant speech effect is the finding that regardless of whether the speech sounds are familiar or unfamiliar (e.g., a language that participants do not speak or nonsense syllables), background speech interferes with an individual’s ability to serially recall visually presented stimuli (Colle & Welsh, 1976; Salamé & Baddeley, 1982). This finding not only suggests that when stimulus presentation is auditory, access to the phonological store is obligatory, but also that the encoding of stimuli entering this store is at the phonological as compared to semantic level (Baddeley, 1997).

The last major experimental paradigm that is used to support the PL argument is the effect that articulatory suppression has on recall performance. To prevent participants rehearsing presented list items, researchers ask them to repeat a syllable or phrase (e.g., *the, the, the*) overtly during list presentation (Baddeley, Lewis & Vallar, 1984; Baddeley et al., 1975). Murray (1967) found a decrease in recall performance when participants were required to articulate an irrelevant word during item presentation. Further, when list items are phonemically similar, articulatory suppression abolishes the effect that similarity has on STM performance for visually presented stimuli, but not auditorily presented stimuli (Baddeley et al., 1984). In comparison, articulatory suppression removes the word length effect regardless of whether the stimuli are presented visually or auditorily, but only when suppression occurs *during* both list presentation and recall (Baddeley et al., 1984). Based on these earlier research findings, the suggestion is that auditory information gains obligatory access into the phonological store, whereas visually presented information needs to be turned into an phonological trace before entering this store. It is assumed that this extra process of recoding visual information into a verbal trace involves the use of subvocal rehearsal (see Baddeley, 1986, 1997, for a more in-depth discussion of the empirical evidence in support of the PL). Therefore, the idea is that the locus of the PSE derives from the phonological store, whereas the word length effect is assumed to arise from the rehearsal mechanism, which is responsible for maintaining decaying information.

1.5. The influence that the lexical status of list items has on short-term memory performance

One area of research that is enjoying a lot of attention of late is the influence that the lexical status of list items has on STM performance. An innovative experiment was designed by Hulme et al. (1991) to investigate whether ISR performance was influenced by an individual's long-term knowledge of phonology. They found that participants were able to recall familiar words more accurately than either Italian words (e.g., *lago*) or nonwords with an English sound (e.g., *maffow*). This has been termed the *lexicality effect* to reflect the performance advantage for words as compared to nonwords or unfamiliar words (Hulme et al., 1991; Roodenrys, Hulme, & Brown, 1993). The term "redintegration" has been used to describe the process by which prior to output, incomplete phonological traces held in STM are filled in or "redintegrated" by phonological representations that are stored in LTM (Brown & Hulme, 1995; see also Schweickert, 1993). According to this view, in comparison to words, STM performance is lower for nonwords because there are no stored representations available to assist in the reconstruction of a partial trace (Hulme et al., 1991; Hulme et al., 1997). Although research into the effect that lexicality has on STM performance has proliferated, most STM modellers (e.g., Brown et al., 2000) believe that it is sufficient for these models to incorporate (at some level), a generic redintegration process to explain these differences. As such, it is difficult to compare STM models based on the explanations that each generates for the lexicality effect.

1.6. The phonological loop revisited

Please see print copy for Figure 1.1

Figure 1.1 A model of the proposed components of the phonological loop that are involved with the storage of phonological STM items and the influence of long-term learning on this STM store (Baddeley et al., 1998).

Based on research that shows a clear performance advantage for words as compared to nonwords, the PL component of working memory has been modified to reflect the relationship and influence of LTM on STM performance (Baddeley, Gathercole & Papagno, 1998). When stimulus presentation is verbal, auditory information is analysed and held in the phonological store (refer Figure 1.1). However, when stimulus presentation is visual, the information enters the phonological store via the articulatory system. This articulatory system is responsible for both subvocal rehearsal and verbal output. The phonological store represents this information in the form of an STM trace. This trace is influenced by the phonological long-term system through the temporary activation of some sort of network or structure. Although this activation is temporary, it is enough to influence representations in LTM, and in turn for LTM to influence STM representations.

The difference between these memory systems lies in the slow (LTM) and fast (STM) weights (see Figure 1.1). As the modification of LTM depends on slow weights, the recall of new information from LTM may require substantial learning. This can be contrasted with the recall of items from STM that rely on fast weights. Hence, the recall of an item from STM is not dependent on the long-term learning of the item yet is still influenced by an individual's long-term knowledge of a language.

Although the PL component of working memory can account for some robust STM findings, neuropsychological studies suggest that individuals with deficits in phonological STM can still comprehend language (Vallar & Shallice, 1990) and produce speech (Shallice & Butterworth, 1977). These findings have raised questions about the evolutionary need for such a capacity. This has lead some researchers to suggest that the phonological store plays a crucial role in new word learning (e.g., Cowan & Kail, 1996; Gathercole, Hitch, Service & Martin, 1997; but see Snowling, Chiat & Hulme, 1991, for an alternative account). Hence, the current view is that the PL makes vocabulary acquisition possible (Baddeley et al., 1998).

1.7. Conclusion and brief outline of the thesis

As a model of STM, the PL has been highly influential. However, attempts to use this model as a theoretical construct to guide current research, have found it sorely lacking in specificity. For instance, the model does not include mechanisms designed to serially order stimulus input (Houghton, Hartley & Glasspool, 1996). Therefore, this model cannot tell us anything about the different types of errors that occur when performing STM tasks (Conrad, 1965; Ellis, 1980; Healy, 1974). Furthermore, this model does not address how learned items may be represented at the phonological level (Houghton et al., 1996). Finally, it would be impossible to specify whether words are processed into phonemes, morphemes or syllables, let alone predict how these constituents may influence performance when the experimental stimuli are nonwords.

Based on the abovementioned problems with earlier attempts to use a theoretical model to guide research, the current thesis will evaluate STM research findings into the PSE, in light of current STM models. As such, chapter two will be dedicated to describing extant STM models in detail. Chapter three will provide a critical examination of the research findings into the PSE, with a particular focus on evaluating the contradictory results observed in the STM literature. The remainder of the thesis will outline the current research that has been designed to investigate the effect that phonological similarity has on STM performance with words (Chapter 4) and nonwords (Chapter 5) using the standard ISR task. To investigate the effect that overt speech production has on STM performance, the PSE will also be examined using the serial recognition paradigm with both words and nonwords (Chapter 6). Furthermore, the research findings will be summarised and conclusions drawn (Chapter 7). Finally, the implications of the current findings for STM models will be discussed and avenues that should be fruitful for future STM research proposed (Chapter 8).

2. EXTANT SHORT-TERM MEMORY MODELS

2.1. Introduction

Despite the success of the PL component of Baddeley and Hitch's (1974) working memory model, it failed to address an important aspect of everyday human performance: how behaviour is serially ordered (Burgess & Hitch, 1996). The ability to temporally sequence behaviour is an important everyday activity (Brown et al., 2000). For instance, a minimum requirement for interpreting what is meant by an utterance is that an individual must be able to sequentially order individual words that were verbalised (Lashley, 1951). Therefore, it is important that STM models include a mechanism capable of serially ordering presented list items.

One of the earliest models proposed to account for serial ordered behaviour was called the *ordered slot model* (e.g., Conrad, 1965). According to ordered slot models, as each item in a list is presented, it is placed into a box or ordered slot (Conrad, 1965). Once a list has been presented, recalling those items in the correct order from STM involves successively searching through each of the boxes. However, research findings pose a number of problems for these earlier models of serial order. For instance, ordered slot models do not store item and order information separately (Brown et al., 2000). Hence, they cannot explain research findings that suggest that experimentally manipulating the stimuli presented to participants influences item and order memory in distinct ways (Healy, 1974; Wickelgren, 1965d). Also, these models do not specify the mechanism necessary for serially searching through these bins. This enigma has been coined the *reinstatement problem* (Brown et al., 2000). This is an important point in that these types of models were proposed to account for serial ordered behaviour. However, they *assume* that this behaviour occurs without specifying *how* it occurs. Therefore, if STM models are to explain serial ordered behaviour, it is critical that these models specify the mechanism by which list items are retrieved.

A standard finding in the STM literature that can be used to evaluate the utility of STM models is the PSE. As suggested by Nairne and Kelley (1999; p.45), "… the phonological similarity effect has achieved the status of a 'benchmark' finding in the immediate memory literature, and most theories of short-term memory include mechanisms that are specifically designed to account for the phenomenon". Although early models of STM cannot adequately explain current research findings (e.g., Baddeley & Hitch, 1974), they form the foundations of extant STM models. To present an exhaustive summary of these models is beyond the scope of the current thesis. However, exemplar models that have incorporated the mechanisms necessary to account for the PSE will be discussed. These extant STM models can be broken into two broad categories: those based on the feature model of immediate memory developed by Nairne (1990a) and models which are derived from Baddeley and Hitch's (1974) original working memory model.

Research based on the Baddeley and Hitch (1974) model has spawned a plethora of new STM models that have incorporated different mechanisms to account for an array of STM research findings. Two of the most influential of these models in recent times is a mathematical model developed by Brown et al. (2000), called the *oscillator-based associative recall* (OSCAR), and Gupta and MacWhinney's (1997) *verbal STM model*. Further, Burgess and Hitch (1992; 1999) have extended the original working memory model to address a more diverse range of STM findings. To gain a thorough understanding of how each of these models attempts to explain research findings into the PSE, a brief outline of the mechanisms associated with each model is imperative. Essentially, the focus of chapter two will be on outlining the mechanisms behind each model, with a particular emphasis on how each model deals with the problem of serially ordered behaviour. How each differs or builds on other models and the explanatory limitations of each model for current STM research findings will also be described. However, a discussion of how each model explains the PSE will be deferred until all of the current research findings into the PSE have been examined.

2.2. Description of Nairne's feature model of immediate memory

Nairne (1988) originally proposed the *feature model of immediate memory* to explain the recency effect. This is the finding that recall is better for items that occur toward the end as compared to items that occur in the middle few positions of an experimental list (Crowder, 1972). The size of the recency effect is influenced by the modality in which items are presented, such that this effect is larger when items are presented verbally as compared to visual presentation (Conrad & Hull, 1968; Corballis, 1966; Craik, 1969; Murdock & Walker, 1969; Murray, 1966). As with all STM models, Nairne (1990a) has based the feature model on previous research and theory and has extended these ideas to account for a wider range of new memory findings.

Based on the work of William James (1890), Nairne (1990a) has distinguished primary from secondary memory processes. According to Nairne (1990a), interference from previously presented items is the vehicle in which primary memory trace degradation occurs. Primary memory traces are thought to be active representations of presented list items that contain order information in real-time. In comparison, secondary memory is thought of as a more permanent storehouse for an individual's experiences. Regardless of whether the to-be-recalled traces are held in primary or secondary memory, the feature model (Nairne, 1990a) represents these traces as vectors of features that differ in both type and value.

2.2.1. Trace features

According to Nairne (1990a), item encoding into primary and secondary memory occurs simultaneously. A trace can be classified in two ways: traces have features that are *modality-independent* and *modality-dependent*. The term 'modality-independent' refers to the features of list items that are not dependent on the modality in which they were presented. The 'modality-independent' features of a trace are assumed to be speech based (Conrad, 1964), although other representational formats such as semantics are possible. In contrast, the term modality-dependent refers to trace features that are specific to the modality in which the list items were presented (Nairne, 1988). More generally, however, these traces consist of both extra-item (e.g., room cues) and intra-item (e.g., stimulus presentation modality, physical features such as the sound of the voice in which the items were presented or font) characteristics (Nairne, 1990a).

2.2.2. Forgetting according to the feature model

Forgetting occurs in the feature model when adjacent items in an experimental list held in primary memory interfere with each other (Nairne, 1990a). The mechanism responsible for recall errors is assumed to be trace interference as opposed to trace decay. Termed *feature overwriting* (Nairne, 1988), this process is assumed to be trace specific. For instance, a one-to-one relationship holds, such that modality-independent features are restricted in that they can only overwrite other modality-independent features and the same applies to the modality-dependent features of a trace. The extent to which an item is overwritten by a subsequently presented list item is dependent on the similarity of the items (Broadbent & Broadbent, 1981). However, even though similarity is necessary for overwriting to occur, Nairne (1988) suggests that it is not sufficient. A participant must subjectively categorise the presented list items as members of the same group and must be able to use this group membership as a cue during recall (Nairne, 1988).

According to the model, evaluations of group membership are subjective. It is the *participant's* appraisal of whether the presented items belong in the same group that determines the extent to which previously presented list items are overwritten by subsequently presented list items. Support for the subjective nature of a participant's evaluations of group membership comes from the finding that group membership can be manipulated experimentally. Examples include, grouping specific items in presented lists together so that they are temporally separate from other grouped items in the same list (Frankish, 1985; Ryan, 1969a) and manipulating semantic category membership (Nairne, 1990b).

2.2.3. Mechanisms responsible for serially ordering list items

According to the feature model, after a list has been presented, a participant is left with a degraded primary memory trace for each item (Nairne, 1990a). Accessing or recalling these traces is dependent on two things: the *distinctiveness* of the experimental items and their *salience* (Nairne, 1988). This model assumes that these item traces are encoded in the order in which they were presented. Although not specified in detail, Nairne (1988) suggests that the temporal ordering of traces is established using a mechanism such as Estes' (1972) perturbation model of positional coding. Therefore, to gain an understanding of the mechanisms that Nairne (1990a) has incorporated into this model, it is important to outline the mechanisms proposed by Estes (1972) to account for serially ordered behaviour.

2.2.3.1. Distinctiveness Models

Distinctiveness models were proposed to account for the finding that the more distinct or extreme a particular item is, the more likely it is to be recalled (Murdock, 1960). It follows that those items that are further away from each other in a list will be easier to distinguish along a positional dimension (Murdock, 1974). Similarly, those items that occur in the initial and final positions in a list will have fewer neighbours than items that occur in the middle positions. Hence, these types of models can be use to explain the primacy and recency effects found when performing an ISR task (e.g., Johnson, 1991). Although Murdock (1960) suggests that it is an item's position in a list that influences item distinctiveness, Neath (1993a, 1993b) has recently suggested that it is the *temporal dimension* as compared to the *positional distinctiveness* of an item which is important for retrieval.

Distinctiveness models also differ as to whether memory performance is influenced by '*local*' or '*global*' distinctiveness. For instance, recall performance depends on the phonological distinctiveness of a particular item in relation to its closest neighbours in *local* models (e.g., Brown et al., 2000; Burgess & Hitch, 1992; Nairne, 1990a). *Global* models, however, suggest that what makes an item distinctive is the distance that a particular item is away from other items in a list along either a *positional* (Murdock, 1960) or *temporal* (Neath, 1993a) dimension. As Brown et al. (2000) suggest, a limitation of pure distinctiveness models is that they do not specify any sort of mechanism by which list items can be sequentially recalled. As such, distinctiveness models are severely limited in that they can only predict *relative* performance for list items based on an item's serial position within a list. Attempts to rectify this limitation have resulted in the combining of current distinctiveness models with perturbation models (e.g., Nairne, Neath, Serra & Byun, 1997).

2.2.3.2. Perturbation and ordered models

According to Estes' *perturbation model* (1972, 1985, 1997; see also Lee, 1992; Lee & Estes, 1981), list items are associated to control elements as opposed to subsequent list items. These control elements can be thought of as representing the *context* in which the items were presented to the model. According to current perturbation models (Estes, 1997), the retrieval of a particular item depends on whether the item-context node to which it is associated can be reactivated. This model assumes that the item-context connections are reactivated via a cyclic process that rotates through the items in the order in which they were presented. Thus, unlike chaining models, where each item is associated to subsequently presented list items, in these types of models (e.g., Page & Norris, 1998), each item is associated to a context node. As such, it is the nodes, as opposed to the items themselves that are associated with each other.

Please see print copy for Figure 2.1

Figure 2.1 A representation of a hierarchical network model of the memory for chunked items in a list and list items in an experimental session (Estes, 1985).

According to perturbation models, item representations in memory are organised hierarchically (refer Figure 2.1). If the input that an individual hears does not sound like one continuous stream, but rather a few separate or discrete groups, then those utterances will be chunked together to form a group (Estes, 1985). Information about an item within a chunk, a chunk within a list and a list within an experimental session are associated with their own control elements or context nodes that encode information about the *temporal* position of the chunks within lists.

Estes (1985) suggests that items in memory can be retrieved in two ways: by either cycling through or reactivating the item-context nodes sequentially, or by the use of a retrieval cue which is associated to a higher order context or control element. Also, the retrieval of items in memory depends not only on the *temporal* relationship between items, but on an item's similarity or distinctiveness in comparison to other presented list items. However, these models do not specify a mechanism that is responsible for serially cycling through or reactivating the item-context nodes. As such, these types of models suffer from the same sort of *reinstatement problem* that ordered slot models were previously criticised for.

2.2.4. List recall

When the task is to serially recall presented list items, Nairne (1990a) suggests that there are two steps that a participant must complete before each item can be recalled. The first step is that a participant needs to distinguish between list traces and residual modality-independent traces. Secondly, a participant is required to compare the degraded trace held in primary memory to similar traces held in secondary memory. As such, recall is based on the matching of a primary trace to a set of traces held in an individual's secondary memory. This is achieved by matching on similarity, or the number of shared features that the two types of traces overlap on (Hintzman, 1986; Nosofsky, 1986). The *"secondary memory search set"* consists only of items that were presented in the currently active list (Nairne, 1990a). According to Nairne (1990a), the selection of items from the secondary search set is based on the ratio rule. The ratio rule suggests that the chances of recalling a presented item is relative to the similarity of all of the traces to the target item (Gillund & Shiffrin, 1984; Luce, 1959; Nosofsky, 1986).

In other words, "the likelihood of correctly sampling an item will be greater whenever its corresponding primary memory vector retains features that are distinctive relative to other items in the list" (Nairne, 1990a; p.254).

Recently, Nairne (1991; see also Nairne & Kelley, 1999; Nairne & Neumann, 1993) has suggested that participants encode items along a multidimensional memory space. This memory space has a *list* and a *within-list dimension.* The list dimension refers to the discriminability of a particular list in relation to other lists that have been presented. For instance, if a participant's task is to reconstruct five lists after the last list has been presented, a particular list will be more discriminable from other lists if each to-be-reconstructed list consists of items from unique semantic categories. However, if each list is composed of items that belong to the same semantic category, then a particular item's list membership will harder to discriminate. In contrast, the within-list dimension refers to the discriminability of particular items within a list. For example, the ability to discriminate a particular item from other list items will be harder when the stimuli in a list are phonemically similar.

2.2.5. Limitations of the feature model

There are a number of key STM findings that the feature model is unable to provide an adequate explanation for. Firstly, this model cannot account for the word length effect. The word length effect has been used to support the idea that memory is influenced by trace decay which varies as a function of time, as opposed to trace overwriting which is influenced by the similarity of list items (although Neath and Nairne (1995) have recently extended the feature model of immediate memory to provide an explanation for the word-length effect). Secondly, this model does not include a mechanism capable of learning. Hence, it cannot explain the *Hebb repetition effect*, which is the finding that when a list is repeated (i.e., same order of items) a number of times within the same experimental session, performance increases as the number of trials increases (Hebb, 1961).

Although earlier conceptualisation's of the feature model (Nairne, 1988, 1990a), hinted at the possibility that a mechanism such as the one proposed by Estes (1972, 1985; see also Lee, 1992; Lee & Estes, 1981) could be used to serially order behaviour, it was not incorporated into the feature model. More recent additions to the feature model (i.e., Nairne et al., 1997) have combined distinctiveness models with Estes' (1985) perturbation model of serial ordered behaviour. However, as suggested previously, perturbation models are still no closer to solving the *reinstatement problem* (Brown et al., 2000). Thus, these models once again *assume* that serial ordered behaviour occurs without specifying in detail *how* it occurs. As Brown et al. (2000; p.135) suggest,

> *"...any model that simply assumes that the states of the learning-context signal can be reinstated in the correct order at retrieval, without specifying in detail exactly how this occurs, has simply postponed the problem from one of recalling a list of items in the correct sequence to one of recalling a list of learning-context signals in the correct sequence".*

2.3. A network model of the articulatory loop (Burgess & Hitch, 1992)

The connectionist model developed by Burgess & Hitch (1992) was designed to provide a computational framework for the articulatory loop component of Baddeley and Hitch's (1974) working memory model. Hence, this model can account for the same STM findings (e.g., the word length effect, the PSE and the influence that articulatory suppression has on recall when stimuli are presented visually as compared to verbal presentation) as the original working memory model. Burgess and Hitch (1992) have also extended this model to account for some of the research findings that the Baddeley and Hitch (1974) model could not explain. For example, the finding that when an error occurs during recall, a large proportion of these are the transposition of adjacent items (Bjork & Healy, 1974; Healy, 1974). Therefore, the current model has incorporated a mechanism that can maintain the serial order of presented list items.

2.3.1. Outline of the proposed mechanisms

As depicted in Figure 2.2, there are four layers of nodes that represent information locally. The activation level of these nodes has been set between ± 1. Information in this model is only passed forward via weighted connections to other layers when the activation level of a particular node is positive. The activation level of the other nodes that a particular node is attached to, determines the sum of a particular node's activity.

Please see print copy for Figure 2.2

Figure 2.2 An outline of the connectionist model of the articulatory loop (Burgess & Hitch, 1992).

2.3.1.1. Context and phoneme nodes

Information enters this model via both the input phoneme nodes that represent the phonemic make up of individual items and the 'context' nodes which are responsible for associating the context to a particular item (refer Figure 2.2). Although the 'context' and phoneme nodes are assumed to have similar characteristics, the context nodes are used to represent non-phonological information such as the *temporal* order in which the items are presented. Burgess and Hitch (1992) describe the activation of the context nodes as being analogous to a ‘moving window’ in which the context nodes vary as each item is presented. The connections between the context-word and phoneme-word nodes are ‘temporarily’ weighted. These temporary weights are ‘learned’ when list items are presented to the model. The term ‘learning’ according to Burgess and Hitch is used in the classical sense, in that learning occurs by a ‘one shot’ Hebbian adjustment of weights (Gupta & MacWhinney, 1997). In other words, when nodes are active at the same time, connections between the nodes are strengthened. Finally, Burgess and Hitch suggest that because of their temporary nature, both of these connections include a random element or ‘noise’, and decay with the passage of time.

2.3.1.2. Word nodes, output nodes and the competitive filter

Word nodes are excited by the input they receive from both the phoneme and context nodes. The input that a word node receives from the phoneme node consists of both temporary connections and pre-learned permanent connections. The most active item is selected during either overt recall or rehearsal through the use of a competitive queuing (CQ) mechanism (refer Figure 2.3).

CQ is a parallel model of serial order that is based on response competition that is temporally modulated (refer Figure 2.3). According to Burgess and Hitch (1992), there are excitatory connections between the word nodes and the competitive filter and inhibitory connections feeding back to the word node from this filter (refer Figure 2.3). The job of these inhibitory connections is to select the most active word node and then suppress the selected word before the next word is recalled. This is achieved by lateral inhibition. All of these connections are hard wired which means that learning does not occur at this level in the model. Once a word is selected, excitation from the competitive filter is passed onto the output phoneme nodes (refer Figure 2.2).

Please see print copy for Figure 2.3

Figure 2.3 The basic architecture of models that have incorporated a competitive queuing (CQ) mechanism (Houghton et al., 1996).

2.3.2. List recall

The Burgess and Hitch (1992) model treats overt recall and rehearsal in the same way. Once a particular item finds its way to the output node, one of two things can happen: either the item is recalled or the phonemic output for the particular item is fed back into the model via excitatory connections between the output and input nodes (refer Figure 2.2). These feedback connections store item-to-item associations or links. As such, they are 'temporary' weights. Their purpose is to learn the association between the representation of a word to the output node and the representation of a subsequent list item to the input node. It is in this way that the current model forms an articulatory loop.

2.3.3. Mechanisms responsible for serially ordering list items

As suggested previously, the problem of how humans serially order behaviour has a long history in cognition (e.g., Lashley, 1951). Plausible accounts of the mechanisms that may be responsible for this behaviour are only now being explicitly stated. This detail is essential in that not only can these mechanisms be used to explain current research findings, but also so that future research can be guided by the predictions that each model generates for particular STM effects. Therefore, it is essential that a critical review of the Burgess and Hitch (1992) STM model explicitly outlines the mechanisms that have been proposed to account for serially ordered behaviour.

2.3.3.1. Competitive queuing

The CQ (refer Figure 2.3) architecture was designed to account for the sequential memory of items and since its conception, has been extensively used in other STM models (Burgess, 1995; Burgess & Hitch, 1992; Hartley & Houghton, 1996; Houghton, 1990; Houghton et al., 1996; Houghton, Glasspool & Shallice, 1994). The CQ mechanism can be used to account for experimental observations such as the finding that, when a list does not contain any repeated items and a recall error occurs, participants rarely repeat an item they have already recalled (Conrad, 1965). It can also be used to account for the Ranschburg effect, which is the finding that when a list contains an item that has been presented twice, participants generally omit recalling the item a second time (Jahnke, 1969).

According to Burgess and Hitch (1992), models of STM that have incorporated a CQ mechanism are only able to account for how serial order is *retained* in STM by using *relative* activation levels. This limitation derives from the fact that CQ was originally developed to model speech production from LTM (Glasspool, 1995). As such, it has only been implemented at the *recall*, as opposed to both the *recall* and *presentation* levels. As Burgess & Hitch (1999) suggest, the types of errors that occur when recalling lists of stimuli suggest that item activations are influenced by both temporal and phonological influences, as well as CQ among them. Hence, Burgess and Hitch (1992) simulated two different serial ordering mechanisms (i.e., associative chaining and a repeatable context-timing signal) in an attempt to provide a model of STM that included the mechanism necessary to account for the serial ordering of information at both the *presentation* and *recall* levels.

2.3.3.2. Associative chaining

One of the oldest approaches designed to account for serial ordered behaviour is *chaining* (Ebbinghaus, 1890/1964). Generally, associative chaining models of serial order assume that items are stored in memory as a number of pairwise associations (Murdock, 1983). For instance, according to the theory of distributed associative memory (TODAM, Lewandowsky & Murdock, 1989), each item in a list can be used to cue the retrieval of successive list items. Thus, in the sequence *ABCDE*, each list item is associated to the item that was presented in the adjacent position (i.e., *A-B, B-C, C-D, D-E*), hence the term *item-to-item associations* (Burgess & Hitch, 1992). In the Burgess and Hitch model (1992; see also Page & Norris, 1998), there are connections between the output and input nodes such that each item at output can be used as a cue in which to recall the subsequent list item (refer Figure 2.2). Therefore, it is at this level in the model that item-to-item associations are formed.

2.3.3.3. Context-timing signal

Models of STM have also turned their attention to incorporating a context-timing signal in an endeavour to solve the *reinstatement* puzzle (e.g., Burgess & Hitch, 1992, 1996, 1999; Hartley & Houghton, 1996; Henson, 1998). Basic to all models that incorporate context-timing signals, activation gradients or position-item associations, is the assumption that the activation level of nodes corresponding to individual list items are temporally graded: The most active item node at any one time is the item that was in that particular *position* during list presentation. Recall according to these models, is due to both activation-based competition between list items and the feedback of inhibitory connections to other list items (e.g., Burgess, 1995; Burgess & Hitch, 1992, 1999;

Hartley & Houghton, 1996; Henson, 1998; Houghton, 1990; Houghton & Hartley, 1996; Page & Norris, 1998). Models that have incorporated activation-gradients to explain memory for serial order differ in the locus of these gradients. For instance, Page and Norris (1998) assume that these activation gradients are established when list recall begins, whereas Burgess and Hitch (1992; see also Houghton, 1990) assume that this process occurs throughout recall dynamically, in that the activation gradients change as each item is recalled.

2.3.3.4. Which serial ordering mechanism?

Burgess and Hitch (1992) modelled both serial ordering mechanisms separately so that the contribution that each makes to serially ordered behaviour could be assessed independently. As they suggest, "… simulations with little or no chaining come closest to reproducing human behavior…" (Burgess & Hitch, 1992; p.456). For example, chaining models would predict that if the list *A B Á D* is presented to participants and an order error occurs, the sequence *A D Á B* is more likely to be recalled than *A Á B D* (Brown et al., 2000). According to chaining models, these types of errors are more likely to occur because the cues (i.e., *A* and *Á*) used to retrieve items *B* and *D* are similar. However, experimental research findings are clearly inconsistent with this prediction (see Baddeley, 1968; Henson, Norris, Page & Baddeley, 1996). Therefore, according to Burgess and Hitch (1992), STM models need to include both a phonological store to account for the influence that phonemic similarity has on recall performance (Baddeley, 1966), and a time-varying signal to explain serial order errors (Bjork & Healy, 1974). Therefore, "… the core postulate of the model is that the characteristics of short-term memory for serially ordered items arise from the way that timing and phonemic information combine to prompt the competitive selection of each item" (Burgess & Hitch, 1996; p.57).

2.3.4. A second bite of the cherry

The Burgess and Hitch (1992) model was the first attempt to provide a computational model of the PL component of working memory. As such, modifications to this initial model were proposed to provide more plausible explanations for a larger array of STM research findings. For instance, in the initial model, context-item and phoneme-item associations were stored in connections that decayed with the passage of time (refer Section 2.3.1.1). Due to the nature of these decaying connections, correlations between items that were temporally well separated in a list allowed items that occurred later in the list to be replaced with those that occurred earlier in the list. Hence, when an order error occurred an unusually large proportion of these were widely separated. This is in direct contrast to experimental results, which suggest that when an order error occurs, it is more likely that adjacent items in a presented list transpose (Bjork & Healy, 1974; Healy, 1974). In order to solve this problem, Burgess and Hitch (1992) modified the context-timing signal so that the activation level of temporally well separated items was graded, such that those items that were adjacent to each other at list presentation were correlated more highly than those items that were temporally well separated.

This slight alteration as described by Burgess and Hitch (1992), allows the modified model to show a recency effect that was absent in the initial model. According to the modified model, the recency effect derives from the context-timing signal in that the timing signal for the initial and last item in a list is more *distinctive* than other list items. For instance, in this model, non-zero correlations between context states are used to distinguish those items that are adjacent or nearby in a list from those items that are temporally further away. Hence, the context states for the initial and last item in a list share a smaller proportion of non-zero correlations as compared to items that are presented in the middle of a list.

2.3.5. Benefits and limitations of the model

There are a number of STM research findings that the Burgess and Hitch (1992) connectionist model of the PL has been able to account for. These include the decline in recall performance when the stimulus lists are around span length (Guildford & Dallenbach, 1925), the serial position curve showing a primacy effect (Crowder, 1972), the word length effect (Baddeley et al., 1975) and also the PSE (Baddeley, 1966). Different types of recall errors that occur when performing an ISR task can also be modelled. For example, the appearance of item and order errors (Bjork & Healy, 1974), which occur mainly when the stimuli are phonemically similar (Wickelgren, 1965d), can be simulated by this model, as can the unusually large number of paired transposition errors (Bjork & Healy, 1974). Finally, this model can also simulate the effect that articulatory suppression has on both span and the PSE (Burgess & Hitch, 1992).

However, there are a number of key STM research findings that cannot be explained by this model. Firstly, it is unable to account for the finding that, when lists consist of alternating phonemically similar and dissimilar items, recall performance takes the shape of a zigzag pattern with better performance for phonemically dissimilar as compared to similar items (Baddeley, 1968). This limitation was bought about by the chaining architecture used in the current model in that the phonemic component of an item was used to cue the recall of the next list item. Secondly, because information entering the model was treated in the same way, regardless of presentation modality (i.e., visual as compared to verbal), this model cannot account for the differential effects that articulatory suppression has on STM performance when the stimuli are presented visually in comparison to verbal presentation (Baddeley et al., 1984).

In addition, the current model has difficulty accounting for the influence that grouping list items has on recall performance (Ryan, 1969a, 1969b). For example, when list items are grouped temporally (*ABC - DEF*), and an order error occurs, it is generally the *B* and the *E* that are recalled in the wrong position (e.g., *AEC - DBF*). Hence, even though *B* and *E* are recalled in the wrong temporal position, all of the items in a list are recalled in the correct within-group position. The Burgess and Hitch (1992) model fails to account for this research finding because the model assumes that the composition of all experimental lists is invariant. Further, although a repeatable context-timing signal was proposed to account for serially ordered behaviour, a plausible mechanism necessary to operate this signal was not specified (Brown et al., 2000). Hence, the Burgess and Hitch (1992) model *assumes* that this behaviour occurs without specifying a biologically plausible mechanism that could account for *how* it occurs. Finally, a learning mechanism was not built into the model. Therefore, it cannot be used to account for STM findings such as the Hebb repetition effect, the lexicality effect or the effect that frequency has on STM performance.

2.4. Modifications to the Burgess and Hitch (1992) model

In recent years a number of key improvements have been made to the original computational model described by Burgess and Hitch (1992). As such, it will be important to critically examine the influence that these modifications have had on the workings, and hence the explanatory power of the current Burgess and Hitch (1999) model of STM.

2.4.1. Temporal grouping explained

Hitch, Burgess, Towse & Culpin (1996) conducted a number of experiments to investigate the influence that temporal grouping has on ISR performance. They manipulated the experimental stimuli in terms of item length, the similarity of the list items and varied the presentation conditions (i.e., visually, verbally or with articulatory suppression). The results of this research have increased our understanding of the mechanism that STM models need to incorporate to explain the effect that grouping has on ISR performance in two major ways. First of all, Hitch et al. showed how a modification to the context-timing signal could account for the temporal grouping effects observed. Secondly, they specified a biologically plausible mechanism that could account for *how* serial ordered behaviour occurs. According to Hitch et al. there are two temporal grouping effects that need to be explained. The first of these is the finding that temporal grouping improves overall recall performance in comparison to when the stimuli are not grouped. Second, is the increase in the prevalence of errors that although the recalled items retain the correct within-group position, they are recalled in the wrong temporal position (Ryan, 1969a, 1969b).

As suggested previously (Burgess & Hitch, 1992), the context signal can be thought of as a moving window, in that the context *changes* as each item is *presented* to reflect the rhythm of an experimental list (refer Figure 2.4). Hence, this model assumes that the moving window is triggered at list presentation. As such, when list items are not grouped, the context-timing signal cycles through the list of items in the order in which they were presented. However, when experimental items are temporally grouped, the pauses between each group disturb the timing of the context signal (Hitch et al., 1996).

This disruption to the timing of the context signal is assumed to be analogous to restarting or reactivating the moving window (refer Figure 2.4). When the context-timing signal is implemented in this way, item recall is higher when the list items are grouped in comparison to when participants are presented with lists where the items are not temporally grouped. Also, the number of within-group positional errors also increase (Hitch et al., 1996). These effects derive from the idea that when there are two sets of timing signals, as is the case when lists are grouped, the overall similarity between the items across a list decreases, whereas the similarity of items which occur in the same within-group position increases (Hitch et al., 1996).

Please see print copy for Figure 2.4

Figure 2.4 The "context" timing signal. The filled circles represent active nodes whereas the unfilled circles are assumed to be inactive nodes. The t is used to represent serial position. (A) The usual pattern of the context signal when the lists are ungrouped as depicted by the set of temporal oscillators (Set 1). (B) The second set of context nodes are supported by a set of temporal oscillators (Set 2) which reset not at the start of recall as in (A), but after each pause (Burgess & Hitch, 1999; Hitch et al., 1996).

2.4.1.1. Mechanism responsible for generating the timing signal

One of the most innovative experimental endeavours of recent times has been research into the mechanism which may influence timing signals in general (Treisman, Cook, Naish & McCrone, 1994; see also Church & Broadbent, 1990, 1992). According to Church and Broadbent (1990), oscillators are hierarchically structured. It is the activation level of these temporal hierarchies that provides the biological mechanism necessary to sequentially order behaviour. Church and Broadbent (1990) suggest that different sets of oscillators deviate around modal frequencies. It is this variability in the frequencies of oscillators, coupled with the synchronisation of two (or more) rhythmical cycles that allows the oscillators to capture the external rhythm of to-be-recalled sets of items. The use of oscillators as a biological mechanism in which sequentially ordered behaviour occurs is supported by research that suggests that frequency oscillations in complex systems, such as the brain, can emerge naturally (e.g., Kauffman, 1993).

2.4.2. Absence of a separate competitive filter node

As suggested previously, the original Burgess and Hitch (1992) model could not simulate the zigzag serial position effect on recall performance when the experimental stimuli alternated between phonemically similar and dissimilar list items (Baddeley, 1968). According to the original model, a particular item's node was activated when an item was presented, as well as other similar items that were presented in the list. As such, the context states became associated to both the correct item and all of the items that were phonemically similar to the target item (Burgess & Hitch, 1996). However, the context node is also responsible for coding the position of items within a list. As such, associations between items that are temporally close to each other in experimental lists are also more highly activated. Accordingly then, when lists of alternating phonemically similar and dissimilar items were presented to the model, these two effects interacted such that when an order error occurred, similar items replaced dissimilar items at recall. This finding has lead Burgess (1995; see also Burgess & Hitch, 1996) to suggest that the mechanisms responsible for serial ordering and phonemic similarity should be more separable.

Furthermore, in the original model (Burgess & Hitch, 1992), the most active response was selected via the competitive filter (refer Figure 2.3). Therefore, not only were phonemically similar items highly activated at the item level but also items that were more temporally similar were also highly activated at this level. Implementing the "winner take all" component directly onto the items as compared to incorporating an extra competitive filter layer means that only the winning item is ever associated to both the phoneme and context nodes.

2.4.2.1. The presentation and recall of items

In the Burgess and Hitch (1992) model, list presentation and recall was modelled in the same way. However, the idea that the serial ordering mechanism and the influence that phonemic similarity has on ISR should be more separable, has lead to differences in the way in which Burgess (1995) has modelled item presentation and recall. For instance, when the model (i.e., Burgess, 1995) is presented with an item, phoneme and context layers are activated simultaneously (refer Figure 2.2). This activation is sent to the item level via both the context-item and phoneme-item connections. Selection of the most active item occurs at the item level via the CQ mechanism. Within the model, a single layer represents input and output nodes. Hence, it is at this point in the model that context-item, phoneme-item and item-phoneme connections are learned and also decay with the passage of time. Once an item has been selected, it is suppressed by the CQ mechanism that has been directly implemented onto the item nodes. As each item is presented to the model, the context signal gets updated. This process is repeated as each item is presented to the model (Burgess & Hitch, 1996).

At recall however, the context signal is reset to reflect the pattern that occurred when the first item was presented to the model (Burgess & Hitch, 1996). This activation spreads to the item nodes via the context-item associations to select the most active item node. The selected item node feeds its activity to the phoneme layer via the item-phoneme connection and back again via the phoneme-item connection. Hence, the most strongly activated item is then selected for output. Note that the activation of context-item, item-phoneme and phoneme-item connections are sequential in this model, thus making the influence that serial order and phonemic similarity have on STM performance more separable. The remainder of this process is identical to when the stimuli are presented to the model with learning, decay, and suppression of the selected items. Finally, the context is then updated and the process starts again until all of the list items have been recalled.

2.4.2.2. Reproducing the zigzag serial position curve

One of the more fine grained experimental findings that STM models have been unable to account for is the finding that when stimulus lists consist of alternating phonemically similar and dissimilar items, recall is lower for similar, as compared to dissimilar items (Baddeley, 1968). This results in a serial position curve that looks like the teeth of a saw. STM models that have modified the CQ mechanism, such that the ordering and phonemic components are more separable, can simulate these findings (Burgess, 1995; Burgess & Hitch, 1996, 1999). For instance, in the original model (Burgess & Hitch, 1992) the phonemic component of a item was used to cue the recall of the next list item. Thus, when an error occurred, it was just as likely to be dissimilar as compared to a similar list item that was transposed (Burgess & Hitch, 1992). However, by modifying the CQ mechanism as described above, the CQ does not operate at the phonemic level and as such, phonemically similar items are no longer queuing the recall of a subsequent list item. Hence, only currently active phoneme nodes and context nodes at each time step are associated with the winning item (Burgess, 1995).

2.4.3. A mechanism responsible for learning

Burgess (1995) assumes that all of the model's connections (i.e., context-item, phoneme-item and item-phoneme) have "slow" and "fast" weights that are presumed to be malleable through the use of a Hebbian adjustment of weights. The slow varying weights are assumed to be responsible for long-term learning effects. In contrast, the fast weights decay with the passage of time and are responsible for short-term learning effects. There are two ways that learning occurs in the model: List presentation order is learned by strengthening the connections between context-timing and item nodes, whereas word pronunciation can be learned by strengthening the phoneme-item connections.

Accordingly then, the Hebb repetition effect (Hebb, 1961) results from the strengthening of the context-item associations by way of the long-term component or slow weights. Hence, recall accuracy should increase when experimental lists are presented more than once. However, serial order intrusions (Conrad, 1960) or the error of recalling an item that was presented in a previous list, in the same position in the current list, should also increase. Further, the word frequency effect on ISR performance is reflected in the strength of the long-term weights that increase each time an item is presented to the model. Finally, the lexicality effect arises from the idea that in comparison to when familiar words are presented to the model, only short-term connections between item-phoneme and phoneme-item nodes are learned during list presentation (Burgess & Hitch, 1996). The model assumes that when the stimuli are nonwords, the selection of item nodes is arbitrary and as such nonwords not benefit from the increased activation of long-term item-phoneme or phoneme-item connections which are unaffected by decay. Hence, the inclusion of varying “slow” and “fast” weights provides the Burgess (1995; see also Burgess & Hitch, 1996) model with the mechanism necessary to account for STM findings such as serial order intrusions (Conrad, 1960), the Hebb repetition (Hebb, 1961), frequency (Watkins, 1977) and lexicality effects (Hulme et al., 1991) on STM performance, without changing the models structure.

2.5. The most current version of the phonological loop model (Burgess & Hitch, 1999)

The current version of the computational model of the PL (Burgess & Hitch, 1999) includes most of the modifications as described above (i.e., except for a single input-output phoneme layer). For instance, the current model has incorporated the modified context-timing signal (Hitch et al., 1996) and specified the biologically plausible mechanism necessary to explain *how* sequential behaviour occurs (Treisman et al., 1994). Also, the current version does not include a separate competitive filter, has kept the mechanisms responsible for serial ordering and phonemic similarity more separable (Burgess, 1995; Burgess & Hitch, 1996) and has incorporated both 'fast' and 'slow' weights to account for the influence of learning on ISR performance (Burgess, 1995). However, as Burgess and Hitch (1999) suggest, a successful STM model of the PL should be able to account for the diverse array of current experimental research findings which are thought to arise from this component of Baddeley's (1986) working memory model. As such, the current exposition of the Burgess and Hitch (1999) model will concentrate specifically on the extensions to the model that were not included in previous accounts (i.e., Burgess, 1995; Burgess & Hitch, 1992, 1996). Specifically, the inclusion of separate input and output phoneme nodes to make the PL model consistent with neuropsychological research findings (Morton & Patterson, 1980; Shallice, McLeod & Lewis, 1985) as well as the influence that presentation modality (Crowder, 1972), articulatory suppression (Baddeley et al., 1984) and recognition (Sternberg, 1969) - as compared to recall - has on STM performance.

2.5.1. Connections between input and output phoneme nodes

The current model of the PL assumes that speech input and speech output are separate subsystems (refer Figure 2.5). For instance, performance is severely hindered when participants are required to simultaneously execute two input or two output tasks (i.e., dual-task paradigm). In contrast, little interference on performance is found when participants are required to simultaneously perform one input and one output task (e.g., Shallice et al., 1985; see Morton & Patterson, 1980 for further arguments for the separation of speech input and output processes). Hence, the current model (i.e., Burgess & Hitch, 1999) has incorporated two phoneme layers, one for speech input and another for speech output (refer Figure 2.5).

Please see print copy for Figure 2.5

Figure 2.5 The basic architecture of the Burgess and Hitch (1999) connectionist model of the articulatory loop.

Burgess and Hitch (1999) have also used neuropsychological research findings to identify how the connections between the input and output phoneme nodes should be implemented into the model. For example, *conduction aphasia* is assumed to result from damage to the neural pathway between Wernicke's (i.e., language comprehension) and Broca's (i.e., speech production) areas (Pinel, 1997) and is said to be independent of "…simple deficits in speech comprehension or production" (Burgess & Hitch, 1999; p.555). Neuropsychological evidence suggests that there are two subclasses of people with conduction aphasia: those individuals that are unable to verbally reproduce a single, long, low frequency word that has just been presented to them (Goodglass & Kaplan, 1972), and those that, although being able to verbally reproduce words, are unable to retain the correct order (Shallice & Butterworth, 1977; Vallar & Baddeley, 1984). Hence, not only have Burgess and Hitch (1999) implemented a pathway from input to output phoneme nodes that is connected via the ordering mechanism and item nodes, but they have included a second direct pathway between the input and output phoneme nodes (refer Figure 2.5). The connections between the input and output phoneme nodes via the ordering and item nodes are assumed to be unidirectional excitatory connections that are temporarily weighted, include a random element or 'noise' and decay with the passage of time. In comparison, the direct pathway is assumed to be a bi-directional hard-wired connection. As such, activation of an output node is automatically reproduced in the input node and vice versa when an input node is activated (refer Figure 2.5).

2.5.1.1. The importance of a bi-directional connection between output and input phonemes

The direct output-input connection and the phonemic feedback that this connection permits, have a number of benefits when incorporated into the model. Firstly, feedback from the output to the input phoneme nodes permits visually presented stimuli entering the system via the item node to activate input phoneme nodes by way of subvocalization. Hence, it also allows subvocal rehearsal more generally (i.e., regardless of presentation modality) to reactivate input phoneme nodes, thus refreshing or relearning list items. Secondly, it permits item selection at output to be influenced by phonemic feedback. Although an oversimplification (as with the Burgess (1995) model), when recalling an item, temporal information is passed on to the item layer. Before an item is selected, information is passed to the output phoneme layer, which is reproduced in the input phoneme layer. From here, phonemic information excites the item layer and the item with the highest activation level is recalled. According to Burgess and Hitch (1999) the phonemic feedback in the model can be thought of in terms of 'hearing one's inner voice' whenever rehearsing or reading a list item.

2.5.2. Presentation modality effects

According to Burgess and Hitch (1999), when the experimental stimuli are presented visually, information enters the model via the item node. From here the information cycles through the loop as usual with item nodes exciting output phonemes, output phonemes to input and then input phonemes to item nodes (refer Figure 2.5). When presented verbally however, information enters the input phoneme nodes directly via an auditory input buffer. The auditory information is maintained in the auditory input buffer until the next item is presented. Therefore, the short-term associations between the input phoneme and item nodes and also the item and output phoneme nodes are maintained for longer for the last list item presented (refer Figure 2.5). Hence, the standard recency effect is found with better performance on the last list item when the stimuli are presented auditorily as compared to visual presentation (Conrad & Hull, 1968; Corballis, 1966; Murray, 1966).

2.5.3. Effects of articulatory suppression

The locus of the effect that articulatory suppression has on recall performance derives from the output phoneme layer (refer Figure 2.5). During recall, the time taken to output items limits the number of items that can be recalled (i.e., the word length effect). According to Burgess and Hitch (1999), articulatory suppression during recall disrupts the relationship between the length of the list items and the rate at which they can be recalled. Hence, the finding that the word length effect is only removed when articulatory suppression is maintained during both list presentation and recall (Baddeley et al., 1984).

Suppression also leads to the activation of irrelevant output phoneme nodes and, hence, causes the activation of irrelevant input phoneme nodes via phonemic feedback. As suggested previously, visual information entering the system passes through the output phoneme nodes before gaining access to the input phoneme nodes. According to the model, the strength of the connections between the input phoneme and item nodes is dependent on the number of phonemes that have been activated in the input phoneme layer (Burgess & Hitch, 1999). As such, the activation of irrelevant input phoneme nodes impedes the activation of visually presented list items. When the stimuli are presented auditorily however, information enters the input phoneme layer directly via the auditory buffer (refer Figure 2.5). Thus, the rehearsal or the relearning of decaying weights is blocked by suppression when list items are presented auditorily. This decreases overall recall performance as compared to when performing an ISR task in the absence of articulatory suppression (Murray, 1968).

2.5.4. Recognition memory

Burgess and Hitch (1999) assume that all of the short-term connections (i.e., input-item, item-output and context-item) - except for the hardwired direct connections between the input-output phoneme nodes - decay with the passage of time. Thus, when participants are required to perform a probed recognition task, the extent to which an item will be remembered as being presented in a particular list, will be dependent upon the extent to which the particular item's representation has decayed. Hence, more recently presented items will have stronger connections between the input-item and item-output nodes. As a consequence, the phonemic feedback that these connections generate should aid in the identification of a probed item that was presented later in a list. This is consistent with research findings that suggest that the earlier a probe's serial position in a list, the higher the number of identification errors (i.e., either incorrectly saying that a probed item was presented in a previous list when it was not or that a probed item was not presented in a previously presented list, when it was) and the longer it takes participants to make a response (McElree & Dosher, 1989; Monsell, 1978).

2.5.5. Benefits and limitations of the current model

Overall, the currently proposed network model accounts for a large majority of STM research findings. As they suggest, the most recent version of the PL "… is capable of explaining a wider range of psychological data on verbal STM than any other current model of this type" (Burgess & Hitch, 1999; p. 577). However, as Burgess and Hitch (1999) admit, there are still a number of limitations that need to be addressed. For instance, earlier research findings suggested that the PSE is abolished when the experimental stimuli are presented visually (e.g., Baddeley et al., 1984). However, when phonemically similar items are presented to this model visually, the PSE persists.

According to Burgess and Hitch (1999) this is a consequence of having long-term connections or slow varying weights between the item-output and input-item nodes when stimuli presentation is visual. However, recently Fallon et al. (1999) have found that when the stimuli are presented visually, the PSE remains under articulatory suppression. Hence, more work needs to be conducted in this area before implementing any modifications to the current model.

A second limitation identified by Burgess and Hitch (1999) is that the current model assumes that item selection occurs at the lexical level. However, research findings suggest that the selection of to-be-recalled information may also operate at both the sub-lexical (Nimmo & Roodenrys, 2002) and supralexical levels (Ericsson & Chase, 1982; Miller, 1956; Simon, 1974). Another related issue concerns the problem of how individual phonemes are ordered for recall. Burgess and Hitch (1999) suggest that a model - such as Hartley and Houghton's (1996) linguistically constrained model of STM - could be incorporated into their model to provide an account of phoneme ordering. This linguistic model has both the mechanism to order information entering the system at a sub-lexical level, and to place language-specific constraints on speech production. Currently, however, these two models have not been merged.

2.6. Description of the oscillator-based associative recall (OSCAR) model (Brown et al., 2000)

The benefit of the current PL model (Burgess & Hitch, 1999) outlined above, is evident by the extensive amount of STM research findings that it is able to account for. However, the value of the *oscillator-based associative recall* (OSCAR) model lies not in its uniqueness as compared to the previously discussed model, but in the depth that Brown et al., (2000) describe the biologically plausible mechanism which is assumed to operate the context-timing signal. As such, OSCAR was primarily developed to describe how the mechanisms necessary to account for serially ordered behaviour could be incorporated into STM models.

Please see print copy for Figure 2.6

Figure 2.6 An outline of the oscillator-based associative recall (OSCAR) model (Brown et al., 2000).

According to Brown et al. (2000) vectors are used to represent items. There are both *item vectors* and *learning-context vectors* (refer Figure 2.6). Sixteen-element vectors are used to represent each item. Item similarity is represented by vector similarity. The learning-context vectors are used to represent the associations between the learning-context of a list of items and the vectors that represent each list item. These vectors are assumed to be intrinsically dynamic. In other words, as each item is added to a list during stimulus presentation, list learning changes to accommodate these new items. Hence, this process is analogous to the timing signals used in other STM models (e.g., Burgess, 1995; Burgess & Hitch, 1992, 1996, 1999; Hartley & Houghton, 1996; Houghton, 1990; Houghton et al., 1994).

2.6.1. The learning-context signal

To adequately deal with the problem of serial order, Brown et al. (2000) have identified the properties that a learning-context signal needs to possess. Firstly, Brown et al. (2000; p.136) assume "…that the state of the learning-context signal at the beginning of list learning can be retrieved directly…". Secondly, to represent the serial ordering of items, the learning-context signal does not repeat. Thirdly, the more temporally distinct learning-context signals are from each other, the more distinct states of the learning-context signal are. Finally, these signals are hierarchical, in that they are able to represent contextual information simultaneously.

2.6.2. The use of a hierarchy of oscillators in which to drive the learning-context signal

Brown et al. (2000) assume that the learning-context signal is driven by an array of oscillators (i.e., 15 different oscillators which combine to make up the 16-element learning-context vectors). Although the learning-context signal consists of both fast and slow moving oscillators that vary at different rates over time, "…it remains constant in overall magnitude or strength" (Brown et al., 2000; p.137). Further, for any one list, the learning-context signal consists of a large number of learning-context vectors. The output from the time-varying oscillators is represented by the symbol Π, and the value of any learning-context vector is dependent upon the Π, at any given moment in time (refer Figure 2.6). It is the variability in the frequencies of the oscillators, coupled with the synchronisation of two (or more) rhythmical cycles that allows the oscillators to capture the external rhythm of to-be-recalled sets of items. Hence, it is this process that makes the serial ordering of list items intrinsically dynamic (see also Burgess, 1995; Burgess & Hitch, 1996, 1999; Hitch et al., 1996). In other words, the sequential behaviour of the learning-context vectors is determined by oscillator outputs and *not* by events that occur in the real world.

However, a note of caution should be exercised when drawing conclusions from the results observed when performance was scored using the item recall criterion. Firstly, item recall performance across all three of the phonemically dissimilar conditions was extremely low. For instance, participants were recalling only 1 to 1.5 items on average out of five. Thus, participants were only able to remember the first (or last) item in each list. This finding makes sense in that when the experimental stimuli are phonemically dissimilar nonwords there is nothing to facilitate the recall of item information (e.g., a shared rhyme ending or repetition priming when the experimental stimuli are phonemically similar). Secondly, when there are high levels of item recall (i.e., as observed in Study 1), item scores produce reliable, meaningful and lawful results.

2.6.3. List presentation and recall

When a list is presented to participants, associations are formed between the list items and the learning-context vectors. These associations are learned each time a new item is added to the list. A Hebbian associative matrix is used to represent the connections between item vectors and the learning-context vectors. The learning-context is intrinsically dynamic, in that the learning-context changes or is constantly updated throughout the presentation of each list item. Hence, the learning-context vector will be different for each list item that is presented. This learning process results in two sets of vector associations: associations between vectors which represent sequential list items and associations between the vectors which represent the difference in the time at which each list item was presented. These connections are referred to as item-context associations and are assumed to be bi-directional (Brown et al., 2000).

To recall a sequence of list items, the states of the learning-context need to be reinstated. The reinstatement of succeeding learning-context states can then be "…used to probe the associative matrix in which the item-context associations are stored" (Brown et al., 2000; p.141). In other words, each of the learning-context vectors is used separately as a cue for the recall of item information. Each vector that is retrieved from memory is compared to both the presented list items and similar items that are stored within the association matrix or learned item-context associations. This is analogous to the redintegration process described by Brown and Hulme (1995; see also Schweickert, 1993). Therefore, according to Brown et al. (2000) the combined activation level of all of the learning-context vectors determines which item is recalled (i.e., the item that is cued most strongly).

2.6.4. OSCAR's benefits and limitations

A difference between the PL model proposed by Burgess and Hitch (1992) and OSCAR, is that the PL model simulated performance when the stimuli were presented auditorily. In comparison, OSCAR simulated human performance when stimulus presentation was purely visual. Also, OSCAR was specifically designed to account for serial ordering effects, and as such, is more limited in scope. In general, however, OSCAR shares a strong likeness to the Burgess and Hitch (1992) model and as such, has retained many of the flaws associated with the earlier PL model (see Section 2.3.5).

However, one of the most innovative and distinguishing features of OSCAR is the implementation of temporal oscillators to provide a mechanism that can account for serially ordered behaviour. This is important for two reasons. Firstly, as suggested previously, the ability to temporally sequence behaviour is an important everyday activity (Brown et al., 2000). Secondly, as STM models have become more detailed, the utility of these models is being gauged by whether the proposed mechanisms can successfully account for serially ordered behaviour (Burgess & Hitch, 1999). Hence, a detailed explanation of the repeatable context-timing signal is necessary if the implementation of this mechanism into STM models is to be successful (see section 2.4 for another STM model which has successfully incorporated this timing mechanism).

2.7. A linguistically constrained model of short-term memory for nonwords (Hartley & Houghton, 1996)

As suggested previously, the Burgess and Hitch (1992, 1999) PL model of STM is limited in that it does not include a mechanism that is capable of serially ordering individual phonemes. Burgess and Hitch (1999) have suggested that a model such as Hartley and Houghton's (1996) *linguistically constrained model of STM for nonwords* could be incorporated into their model to provide an account of phoneme ordering. Hartley and Houghton's (1996) linguistically constrained STM model is not only capable of serially ordering phonemes, but also places language specific constraints of speech production. This model was specifically engineered to bring together models of speech production that are based on linguistic research, with existing models of STM.

2.7.1. Linguistic principles

Hartley and Houghton (1996) have incorporated two linguistic principles into their model: syllable structure and sonority. The first linguistic principle that Hartley and Houghton (1996) have incorporated is that syllables have an internal structure, comprising an onset and a rhyme. The onset consists of the initial consonant or consonant cluster, while the rhyme consists of the vowel and any following consonants (Fudge, 1969; Goswami & Bryant, 1990; Treiman, 1986). The rhyme unit can be further divided into the peak (the vowel) and the coda (the following consonants).

The second linguistic principle that has been incorporated into this model is sonority. Sonority refers to the amount of energy in the speech signal, whereas the sonority principle refers to the fact that in syllables, sonority increases to a peak in the vowel and then decreases. Hence, according to Hartley and Houghton (1996) the strength of the speech trace for the individual phonemes that comprise a syllable differs depending on whether the phoneme is a consonant or vowel. In other words, the strength of the speech trace will not be as strong for consonants as compared to vowels because consonants are shorter in duration and are not as acoustically intense (Hartley & Houghton, 1996).

2.7.2. Description of the model

According to Hartley and Houghton (1996), a minimum requirement needed for STM models to model phonological structure is that they simultaneously represent stimuli at two levels: the phoneme and syllable levels (refer Figure 2.7). To correctly recall a syllable, both the order and the identity of individual phonemes must be remembered. As such, the Hartley and Houghton (1996) model represents stimuli at the phoneme level. One node is used to represent each phoneme at this level in the model. Each syllable at the syllable level is represented by two separate nodes that correspond to the onset and the rhyme. Hence, it is at this level that Hartley and Houghton (1996) impose an internal onset-rhyme structure on syllables, which is consistent with current linguistic research findings (Fudge, 1969; Goswami & Bryant, 1990; Treiman, 1986).

Correctly recalling a list of stimuli requires that participants not only remember the order of individual phonemes (phoneme layer), but also the order in which the syllables were presented. Hence, once the individual phonemes have been parsed into syllabic chunks, the syllable level is responsible for maintaining their order in a list (see Figure 2.7).

Please see print copy for Figure 2.7

Figure 2.7 Outline of the architecture of Hartley and Houghton's (1996) linguistically constrained model of STM. The dashed lines depict connections that are free to vary as a consequence of learning, whereas the strength of the connections between the syllable template and phoneme layers are fixed.

There are two pathways that connect the syllable and phoneme levels: the content and the structural pathway (see Figure 2.7). The content pathway is the direct link between a syllable and the phonemes that comprise the syllable. The structural pathway also has connections from the syllable to phoneme levels. However, these connections are via the 'syllable template mechanism'. This syllable template mechanism is based on sonority, which is the second linguistic principle that Hartley and Houghton (1996) have incorporated into their model.

Please see print copy for Figure 2.8

Figure 2.8 The structure of Hartley and Houghton's (1996) cyclical syllabic template.

For ease of illustration, the syllable template depicts only five slots (refer Figure 2.8). When a syllable is presented to the model, each phoneme is matched to a slot sequentially, depending on whether the phoneme is an onset (the initial consonant or consonant cluster - slots 1 and 2), the peak or vowel (slot 3) or the coda (the final consonant or consonant cluster - slots 4 and 5). Hence, the nonword *blint* uses the slots 1, 2, 3, 4 and 5 whereas the nonword *mef* uses slots 1, 3 and 5. Those syllables that conform to the sonority principle require a single cycle whereas those syllables that violate this principle require more than one cycle. Language-specific constraints are also imposed on phoneme order, such that for English words, the /tl/ cluster cannot occur in the initial (i.e., onset) part of a syllable. Therefore, only phonemes that can legally occupy a certain position in a syllable, in a particular language, can be represented (refer Figure 2.8).

2.7.2.1. List presentation and recall

During list presentation, the model is presented with a stream of phonemes. Each phoneme in the stream activates a single node at the phoneme level. Hartley and Houghton (1996) set the activation level of nodes representing vowels higher than for nodes representing consonants to reflect the idea that the speech traces for consonants are shorter in duration and not as acoustically intense as they are for vowels. As suggested previously, activity from the phoneme nodes establishes connections to the syllable level via two pathways simultaneously. The direct pathway connects the individual phonemes with the syllable unit and is responsible for encoding phonemic content. The indirect pathway is via the syllable template nodes. This pathway is responsible for encoding syllabic structure. Hartley and Houghton (1996) assume that the activation of each template node is *context* dependent. In other words, it depends on which slot the previously presented phoneme occupied. For instance, consonants that follow a vowel will activate the post-vocalic consonant slots. As suggested previously, each syllable is represented by two nodes at the syllable layer: one node for the onset and a second node for the rhyme. According to Hartley and Houghton (1996) the two nodes that are used to represent each syllable are never active at the same time. At the syllable level, activation of a pre-vocalic consonant slot at the syllable template level activates the onset node and activation of either the vowel or a post-vocalic consonant activates the rhyme node. Hence, information about the phonemic content and the structure of each syllable that is presented to the model is simultaneously encoded.

At recall, Hartley and Houghton (1996) assume that a competitive cueing mechanisms (see section 2.3.3.1) is responsible for the activation of the syllables at the syllable level. According to Hartley and Houghton (1996), the main aim of the recall process is to recreate the same pattern across the nodes as was activated during item presentation. Accordingly then, an item follows the same route through the structural pathway as suggested above, except that the layers within this model are activated in the reverse order (i.e., from the syllable layer to the phoneme layer via the syllable template). At the same time, the phoneme nodes receive activation from the syllable layer via the content pathway. Hence, "a phoneme node receives input from both the structural and content pathways" and is only activated when it receives input from both of these pathways in combination (Hartley & Houghton, 1996; p.11). The item is then recalled and the process begins again.

2.7.3. Benefits and limitations of the model

There are a number of advantages in merging a linguistically constrained model of speech production with verbal STM models. Firstly, for a participant to correctly recall a presented list item, constraints need to be placed on the way in which the phonemes in an item are ordered. For instance, if these constraints did not exist and a participant was given the nonword *lof* to recall, they would be just as likely to recall the nonword *fol* or *olf*. However, linguistic research suggests that when a recall error occurs, consonants and vowels rarely substitute for each other, regardless of whether the stimuli are words (Brady, Shankweiler & Mann, 1983) or nonwords (Treiman & Danis, 1988). Secondly, the incorporation of a syllable template imposes language-specific constraints on the ordering of phonemes such that for English words, the /mt/ cluster cannot occur in the initial (i.e., onset) part of a syllable (Fudge, 1969). Finally, it provides the mechanism necessary to account for linguistic research on syllable structure which has found that when a recall error occurs, the vowel-consonant (_VC) as compared to the consonant-vowel (CV_) components of consonant-vowel-consonant (CVC) syllables are more likely to be retained (Kessler & Treiman, 1997; Treiman & Danis, 1988).[2]

Although Hartley and Houghton's (1996) linguistically constrained model of STM for nonwords is able to account for linguistic research findings which suggest that STM performance is influenced by linguistic mechanisms that operate at the sub-syllabic level, this emphasis has limited this models ability to explain a key STM research finding. For example, this model was engineered to explain findings from a small body of research on STM for nonwords. As a consequence, it does not distinguish

---

[2] Although simulations using this model does produce recall errors that retain the _VC as compared to CV_ combinations, it does so to a lesser extent than is observed experimentally. Hartley and Houghton (1996; p.17) have suggested a modification to their model, such that "…competition occurs between the onsets and between rhymes, rather than between syllabic units".

between nonsense syllables (e.g., *maf*) and lexical items (e.g., *mat*). Hence, this model is unable to account for the lexicality effect on STM performance - the finding that memory span is higher for words as compared nonwords (Hulme et al., 1991). As suggested by Burgess and Hitch (1999), a possible solution to this problem would be to merge Hartley and Houghton's (1996) linguistically constrained model of STM for nonwords with verbal STM models that have been designed to explain STM research findings from studies that have used words. Gupta and MacWhinney (1997) have produced such a model.

2.8. Origins of Gupta and MacWhinney's (1997) verbal short-term memory model

Gupta and MacWhinney (1997) have developed a STM model to bridge the gap between verbal STM performance and vocabulary acquisition. This model is based on a combination of Hartley and Houghton's (1996) 'linguistically constrained model of STM for nonwords' and the Burgess and Hitch (1992) PL model. Therefore, as with OSCAR (Brown et al., 2000) Gupta and MacWhinney's (1997) model of verbal STM retains many of the flaws associated with this earlier PL model (see Section 2.3.5). However, it does provide insight into the level at which STM models should be modified to include linguistic constraints on STM performance. As such, the description of Gupta and MacWhinney's (1997) verbal STM model that ensues will focus mainly on the level at which they have incorporated Hartley and Houghton's linguistically constrained model of STM for nonwords.

According to Gupta and MacWhinney (1997), the phonological chunk layer represents groups of syllables whereas the phoneme layer represents individual phonemes (refer Figure 2.9). The phonological chunk layer is assumed to have a *topological organisation*. Based on a distributed feature map model of the lexicon (DISLEX; Miikkulainen, 1990), topological organisation is the idea that units have a two-dimensional spatial structure. Originally devised to account for the presentation of visual word forms, Gupta and MacWhinney (1997) have suggested that this type of topological organisation could also be useful for organising phonologically similar items. For example, phonologically similar items (e.g., *dog* and *bog*) represent more similar information and, hence, occupy space that is closer together than dissimilar items (e.g., *dog* and *fat*). In this model both the phonological chunk and phoneme layers have a CQ structure (refer Figure 2.3).

2.8.1. Brief description of the model

Semantic information can be sent to, and can receive input from, the model's phonological chunk layer (refer Figure 2.9). Semantic or lexical information is topologically organised in the same way as the phonological chunk layer, except that it is organised for meaning rather than phonological similarity. The phonological store represents the element to which presented words can be temporally bound in a sequential order and is analogous to the item layer in the Burgess and Hitch (1992) PL model. As with the phonological store, the context maintenance (queue) also encodes the sequences of information entering the chunk layer in a spatial pattern. Whereas encoding in the phonological store is an automatic process, context maintenance encoding is a controlled process. The syllable template is between the phonological chunk layer (i.e., which represents groups of syllables), and the phoneme layer. The elements of the syllable template are sequentially activated (i.e., most active phoneme) and are activated in this model, regardless of whether the stimuli are words or nonwords.

Please see print copy for Figure 2.9

Figure 2.9 Proposed model of verbal STM and vocabulary acquisition (Gupta & MacWhinney, 1997).

2.8.2. Item presentation and recall

According to Gupta and MacWhinney (1997), information enters the model via two simultaneous routes. When the presented item is familiar, the chunk layer representing the particular word is activated directly. For unfamiliar items, however, a new chunk node is activated. Also, each of the phonemes that make up an item are activated in turn. This activation feeds forward to the phonological chunk layer via the syllable template (refer Figure 2.9). Weights between the phonological chunk layer and phoneme nodes are strengthened via a Hebbian adjustment of weights. Once the individual phonemes that make up a syllable are cycled through the syllable template mechanism they are represented by a syllable node at the phonological chunk layer. Hence, unlike the syllable layer in the Hartley and Houghton (1996) where two nodes are used to represent each syllable (i.e., an onset and rhyme node), in this model, one node is used to represent each syllable at the phonological chunk layer. The job of the context maintenance queue is to temporally sequence these syllable nodes into chunks, and the chunks into a sequence. The connections between the semantic layer and phonological chunk layer, and between the phonological store and chunk layer, are strengthened via a Hebbian adjustment of weights. This process is assumed to be automatic and could be likened to the 'slow' and 'fast' weights used in the Burgess and Hitch (1999) model to reflect the influence that LTM has on STM performance.

2.8.3. Benefits and limitations of the model

As outlined previously (see section 2.7.3), there are a number of advantages to merging Hartley and Houghton's (1996) linguistically constrained model of STM for nonwords with verbal STM models. For instance, it is near impossible to see how a STM model could account for research findings that suggest that STM performance is influenced by linguistic mechanisms that operate at the sub-syllabic level, without the incorporation of an additional mechanism that is capable of imposing both syllable structure and language-specific constraints on phoneme order (Brady, Shankweiler & Mann, 1983; Fudge, 1969; Kessler & Treiman, 1997; Treiman & Danis, 1988). However, Gupta and MacWhinney's (1997) verbal STM model, as with OSCAR (Brown et al., 2000), is based on the original computational model of the PL developed by Burgess and Hitch (1992) and as such retains many of the flaws associated with this earlier model (see Section 2.3.5).

Finally, one major difference between Hartley and Houghton's (1996) linguistically constrained model and Gupta and MacWhinney's (1997) verbal STM model is in the level at which each model represents individual syllables. For instance, although the Gupta and MacWhinney (1997) model includes both a phoneme level and a syllable template mechanism, it only uses one node to represent each syllable at the phonological chunk layer. In contrast, the Hartley and Houghton (1996) model uses two nodes (i.e., one node for the onset and another for the rhyme) to represent each syllable at the syllable level. Although this distinction does not seem too important, as Hartley and Houghton (1996) suggest, without representing each syllable in terms of distinct onset-rhyme nodes, this model would show large numbers of pre- and post-vocalic transpositions. This is clearly inconsistent with previous research (e.g, Ellis, 1980) which suggests that when a recall error occurs, the erroneously recalled phoneme tends to retain its syllable position. As such, it is unclear how the Gupta and MacWhinney (1997; p.304) model could provide an "…account for why such transposition errors adhere to the constraints of syllable structure" without representing syllables at a syllable layer in terms of items with a distinct onset-rhyme structure.

2.9. Summary and Qualifications

The PSE has been instrumental in developing theories about how serial recall is accomplished and nearly all current theories[3] have an explanation for how phonological similarity influences the retrieval of serial order. The aim of chapter two was to provide a detailed outline of some models that make use of different mechanisms to account for an array of STM findings, and more specifically, that provide explanations for the influence that phonological similarity has on both item and order retention. For instance, Nairne's (1988) feature model of immediate memory explains STM performance in terms of processes and principles that are applicable to almost any type of information (e.g., spatial locations, object or verbal items). One other STM model that uses general priniciples to explain STM performance is OSCAR (Brown et al., 2000). This STM model is rather unique in that OSCAR describes the biologically plausible mechanism which is assumed to operate the context-timing signal. In other

---

[3] One exception to this is the distributed model of memory for serial order (SOB) developed by Farrell and Lewadowsky (2002), which by their own admission is unable to account for the effect that phonological similarity has on STM performance.

words, Brown et al. (2000) were able to describe how the mechansisms necessary to account for serially ordered behaviour could be incorporated into STM models.

In contrast, the phonological loop model of STM (Burgess & Hitch, 1999) views STM as a specialised language learning device. Other STM models that naturally fall into this category are the primacy model (Page & Norris, 1998) and the start-end model (Henson, 1998). To present an exhaustive summary of thee types of models was beyond the scope of the current thesis. Therefore, the PL model which "…is capable of explaining a wider range of psychological data on verbal STM than any other current model of this type" (Burgess & Hitch, 1999; p. 577) was examined and used as an example to represent other models of this class such as the start-end (Henson, 1998) and primacy (Page & Norris, 1998) models.

Although Hartley and Houghton's (1996) linguisticaly constrained model of STM for nonwords and Gupta and MacWhinney's (1997) verbal STM model can be classified as models that view STM as a specialised language learning device, each was engineered to unify STM research findings and findings from other disciplines that directly impact on the mechanisms that STM models need to incorporate if they are to critically examine the idea that STM is indeed a specialised language learning device. For instance, Hartley and Houghton's (1996) STM model was designed to bring together models of speech production, that are based on linguistic research, with existing models of STM. This model is unique in that it includes a mechanism that is capable of serially ordering individual phonemes. Further, Hartley and Houghton (1996) have also incorporated the linguistic principles of syllable structure and sonority into their model to account for current linguistic research findings (Treiman, 1986). In contrast, Gupta and MacWhinney's (1997) verbal model of STM was developed to bridge the gap between verbal STM performance and vocabulary acquisition. They proposed a conceptual hybrid model by describing how the Hartley and Houghton (1996) model can be incorporated into the Burgess and Hitch (1992) connectionist model of STM and as such provide STM models with mechanisms that can be used to explain how language is acquired.

In summary, chapter 2 examined a number of STM models in detail. These STM models are as diverse in their theoretical orientation (e.g., general vs specific principles), and their level of specificity (e.g., the idea that items are represented at the word as opposed to phoneme or sub-syllabic level) as they are in the mechanisms that each uses to account for serially ordered behaviour. However, for the present purpose, there is one thread that binds these models together: each model provides insight as to how, and what type of information is represented in STM and in doing so examine the relationship between STM representations and the influence that both lexicality and similarity have on STM performance. Both of which are the focus of the current work.

2.10. Where to next?

As has been suggested previously, the worth of an STM model lies in its ability to explain current experimental research findings. One standard finding in the STM literature that can be used to evaluate the utility of STM models is the effect that similarity has on STM performance. As Nairne and Kelley (1999) suggest, nearly all theories or current STM models include some sort of mechanism that has been specifically designed to account for the PSE. However, as the number of studies investigating the PSE increase, so do the smaller number of contradictory findings (e.g., Coltheart, 1993; Gathercole et al., 1982; Watkins et al., 1974). Further, the research question that is currently being asked is not, 'What is the effect of phonological similarity on recall performance?' but 'What effect does operationally defining similarity in different ways have on recall performance?' (Fallon et al., 1999) As such, if current STM models are to guide future research directions they need to be at the level of specificity necessary to account for such fine grained experimental research. Therefore, in chapter three, important experimental research findings into the PSE will be critically examined.

3. EXPERIMENTAL RESEARCH FINDINGS INTO THE EFFECT THAT PHONOLOGICAL SIMILARITY HAS ON SHORT-TERM MEMORY

3.1. Introduction

As has been described previously, the PSE is the finding that performance is worse when the stimuli in a list sound similar to each other as compared to when they do not (Conrad & Hull, 1964). The effect that similarity has on order memory is such a robust finding in the STM literature that, "… most theories of short-term memory include mechanisms that are specifically designed to account for the phenomenon" (Nairne & Kelley, 1999; p.45). Unique experimental paradigms have also been engineered to address different questions about the influence that similarity has on STM. Hence, to gain an understanding of the influence that similarity has on STM performance, it will be necessary to trace the lineage of earlier experimental research findings in this area.

Although early research findings with letters suggested that the PSE on order memory is a stable STM finding, when performance is measured for the recall of item information, the results are contradictory (e.g., Coltheart, 1993; Gathercole et al., 1982; Watkins et al., 1974). Three suggestions have been proposed to account for the discrepant results that have recently been reported in the research literature: the measure of STM performance employed, the size of the stimulus pools used to select the list items, and how similarity has been operationally defined. As such, it will be important to critically examine the influence that these methodological differences have on the research findings observed when the experimental lists are phonemically similar.

3.2. Different experimental paradigms address different questions

Studies investigating the effect that similarity has on STM performance can be grouped into three categories according to what they aim to investigate: overall recall performance, error patterns, or interference effects. Each category addresses a different question with regards to the influence that similarity has on STM (Sperling & Speelman, 1970). For instance, studies grouped into the overall recall performance class were designed to investigate whether there was a difference in recall performance when the experimental stimuli sounded similar as compared to when they did not. However, these earlier studies only suggested that there was a difference. As such, the second class of studies (i.e., error analyses), were designed to investigate *why* this difference occurs. The final class of experiments were concerned with the effect of interference on STM. These types of studies primarily manipulate the acoustic similarity of interfering lists and examine the effect that this manipulation has on STM performance. The research findings across these different experimental paradigms have contributed to our understanding of the nature of STM. As such, a brief discussion of the experimental findings observed using each of these paradigms, and how these results further our understanding of the processes involved in STM will ensue.

3.2.1. The influence of acoustic similarity on short-term memory performance

The first sets of experiments were designed to investigate whether acoustic similarity influenced STM performance. Typically, the dependent measure used to analyse recall performance is the number or proportion of items correctly recalled. Conrad and Hull (1964) found that recall performance was worse when the stimulus lists consisted of letters that sounded similar as compared to distinct sounding letters (see Table 3.1 for some examples of early research findings on the PSE with letters).

Research findings into the PSE using lists of acoustically confusable and non-confusable letters are consistent, in that when scored using the strict scoring criterion (i.e., scored as correct if participant's recalled the correct item in the correct position), STM performance is better for distinct as compared to acoustically similar sounding letters (Baddeley, Lewis & Vallar, 1984; Conrad, Baddeley & Hull, 1966; Conrad & Hull, 1964; Laughery & Pinkus, 1966; Schweickert, Guentert & Hersberger, 1990). This detrimental effect of similarity on STM performance persists, despite the presentation modality (i.e., visual or auditory) or recall method (i.e. written or verbal) employed across different studies (refer Table 3.1). Therefore, although it is clear from earlier research findings that the acoustic confusability of lists of letters impairs STM performance, what is unclear is *why*.

Table 3.1 Experimental Findings on the PSE with Methodological Differences in Stimuli Presentation and Recall Techniques when the Experimental Stimuli were Letters.

	Presentation	**Recall**	**Scoring**
Research			***Strict Recall***
Conrad & Hull (1964)	Visual	Written	Detrimental
Conrad et al. (1966)	Visual	Written	Detrimental
Laughery & Pinkus (1966)	Visual	Written	Detrimental
Schweickert et al. (1990)	Visual	Verbal	Detrimental
Baddeley et al. (1984)	Auditory	Written	Detrimental
Laughery & Pinkus (1966)	Auditory	Written	Detrimental
Sperling & Speelman (1970)	Auditory	Written	Detrimental
Wickelgren (1965d)	Auditory	Written	Detrimental

3.2.2. Types of errors that occur when the stimuli sound similar

In comparison to the above mentioned studies, a second class of experiments were designed to examine the types of errors that participants produce when recalling lists of acoustically confusable items (Conrad, 1962, 1964; Wickelgren, 1965b, 1965c, 1965e, 1966a). According to Conrad (1962), if participants consistently produce the same error for a particular stimulus, this error may reflect a decaying STM trace. In a typical experiment, participants are presented with lists of confusable letters (e.g., *B, C, P, T, V,* or *F, M, N, S, X*). A participant's task is to recall the letters in the order in which they were presented. Conrad (1962, 1964) found that recall errors are not randomly distributed. Rather, similar letters (e.g., *F* and *S*) are confused more often than distinct sounding letters (e.g., *F* and *B*). For instance, when presented with the letter *B* and an intrusion error occurs, a participant is more likely to recall the letter *P*. Further, the error maintains both part of the correct item (i.e., the vowel) and the overall structure of the item (i.e., two phonemes) (Conrad, 1964; Wickelgren, 1965e, 1966a). Furthermore, the effect that acoustic similarity has on the types of intrusion errors that occur is unaffected by whether stimulus presentation is visual (Conrad, 1962) or auditory (Wickelgren, 1965b).

3.2.3. Interference effects when the stimuli are acoustically similar

The last class of experiments are concerned with interference effects on STM performance (Conrad, 1967; Dale, 1964; Dale & Gregory, 1966; Wickelgren, 1965a, 1966b, 1966c). In these types of experiments, a participant is presented with a single item (Wickelgren, 1966b, 1966c), a list of letters (Conrad, 1962; Dale, 1964) or a list of words (Dale & Gregory, 1966). Participants are then presented with a second list that is either acoustically similar or distinct from the original list. Their task is to recall the item or items that were initially presented, in the correct serial order. Although the tasks used to study interference effects on STM vary, when scored using a strict criterion, distractor items that sound similar to a target item interfere with performance to a greater extent than when the distractor items are acoustically distinct (Dale & Gregory, 1966; Wickelgren, 1965a, 1966b, 1966c).

There are a number of other experimental manipulations designed to investigate the interfering effect that similarity has on STM performance. For instance, Conrad (1967) presented participants with lists that varied in the number of acoustically similar stimuli drawn from different classes (e.g., *BCPTV, FSX* and *MN*). Participants were then asked to verbally repeat digits as they were presented, before recalling the letters. Hence, the stimuli employed in the filler task were unrelated to the acoustically similar lists. Using this technique, Conrad found that although errors were more common at longer as compared to shorter retention intervals, the errors that occur are less likely to be acoustically similar to the target item. Based on these findings, Conrad suggested that increasing the retention interval increases item decay and hence, increases the occurrence of errors that are acoustically unrelated to the target item.

3.2.4. Summary

In summary, these earlier studies investigating the effect that acoustic similarity has on STM performance, tell us a great deal about the nature of STM. For instance, the finding of an increase in STM performance when letters are acoustically non-confusable as compared to confusable letters (Conrad & Hull, 1964), suggests that the STM trace is acoustically based. Further, the finding that similar letters are confused more often than distinct sounding letters, suggests that similar items are coded in STM in an analogous way. This finding also suggests that forgetting in STM is not all or nothing. Rather the idea is that partial memory for an item can aid in the retrieval of an item. Lastly, based on the interference studies, Conrad (1964; p. 80) suggests that, "… the more chance there is of acoustic confusion within the stimulus set, the poorer will recall be. It would follow that memory span would be a function of the acoustic similarity of the members of the set". The idea that as similarity increases order memory should decrease persists today as a major assumption within most STM models (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Nairne, 1990a).

Earlier research findings suggest that the effect that similarity has on STM is a robust finding. However, this view has been challenged by contradictory research findings from studies with words or one-syllable nonsense words, as compared to letters (Coltheart, 1993; Gathercole et al., 1982; Poirier & Saint-Aubin, 1996; Watkins et al., 1974). One suggestion proposed to account for these discrepant findings is that differential results emerge as a consequence of the scoring criteria used to analyse performance (Poirier & Saint-Aubin, 1996). Alternatively, the size of the stimulus pools used to select the lists and how phonemic similarity has been operationally defined have also been proposed to account for the contradictory research findings (e.g., Coltheart, 1993; Fallon et al., 1999). As such, the influence that these methodological differences have on current research findings into the PSE will be critically examined.

3.3. Experimental findings into the phonological similarity effect using words

One of the first experiments to investigate the effect that similarity has on recall performance with words was conducted by Baddeley (1966). Participants were presented with lists of words that were either acoustically similar or distinct and their task was to recall the five-word sequences in the order in which they were presented. Consistent with research findings from studies with letters (e.g., Conrad & Hull, 1964; Wickelgren, 1965d), Baddeley (1966) found a detrimental effect of similarity on STM, such that performance was better for dissimilar as compared to similar lists of words (see Table 3.2).

Table 3.2 Experimental Findings into the PSE when the Stimuli are Words and a Strict Scoring Criterion is used to Measure Recall Performance as a Function of Presentation Modality and Recall Method.

Research	Presentation	Recall Technique	Strict Recall
Baddeley (1966; Experiments 1 & 2)	Auditory	Written	Detrimental
Baddeley et al. (1984)	Auditory	Written	Detrimental
Gathercole et al. (1982)	Auditory	Written	Detrimental
Baddeley (1966; Experiments 1 & 2)	Visual	Written	Detrimental
Coltheart (1993)	Visual	Written	Detrimental
Fallon et al. (1999; Experiment 1)	Visual	Written	No Difference
Fallon et al. (1999; Experiment 2)	Visual	Written	Detrimental
Gathercole et al. (1982)	Visual	Written	Detrimental
Poirier & Saint-Aubin (1996)	Visual	Written	Detrimental
Watkins et al. (1974)	Visual	Written	Detrimental
Lian et al. (2001)	Auditory	Verbal	Detrimental
Li, Schweickert & Gandour (2000)	Visual	Verbal	Detrimental

To control for the influence that using different stimuli across conditions may have on recall performance, Gathercole et al. (1982) constructed similar (i.e., rhyming) and dissimilar lists using the same stimulus pool. When scored using the strict scoring criterion, performance was worse when the stimuli rhymed as compared to dissimilar lists (see Table 3.2). Therefore, regardless of whether the same or different stimuli are employed across experimental conditions, the detrimental effect that similarity has on STM performance remains (Gathercole et al., 1982).

In comparison to the previously mentioned studies, Fallon et al. (1999; Experiment 1) found no difference in order memory between similar (i.e., rhyming) and dissimilar six-word lists (see Table 3.2). However, participants were only presented with eight lists per condition. Also, the means were in the direction that one would expect, with better performance for dissimilar as compared to the phonemically similar lists. Hence, the null effect observed in the Fallon et al. study may be due to a lack of power as compared to a true memorial effect. Overall, when scored using a strict criterion, differences in the stimulus sets used across studies, the types of stimuli used (i.e., words or letters), the modality in which stimuli are presented (i.e., visual or auditory), and the recall technique employed (written or auditory), yield the same detrimental PSE on STM performance (e.g., Baddeley, 1966; Li et al., 2000; Lian et al., 2001).

3.4. How important is the scoring criterion?

Until this point, the detrimental effect that similarity has on performance is a robust finding in the STM literature. To further investigate this STM effect, Wickelgren (1965d) examined whether similarity also influenced the recall of item information. Wickelgren measured performance in terms of both order (i.e., strict scoring criterion) and item recall (i.e., scored as correct if a participant recalled an item presented in a given list, regardless of position). When scored using a strict criterion, the findings were consistent with previous research (e.g., Conrad & Hull, 1964), with better performance for distinct as compared to acoustically confusable lists of letters. However, when measured at the item recall level, no differences in the recall of item information were observed (see Table 3.3). These findings suggest that similarity influences the *order* in which items are recalled, rather than the *number* of items recalled (Wickelgren, 1965d).

However, a problem noted by Wickelgren (1965d; p.570) is that the strict scoring criterion "…reflects the combined operation of the recall of items and the recall of the correct position for items". For instance, no order errors are possible if a participant does not recall any of the list items when scored using the item recall measure. In contrast, if a participant recalls all of the presented list items when using the item recall measure, an individual's odds of committing an order error for that particular list have increased. Thus, Wickelgren obtained a measure of order accuracy (i.e., subtracting the number of errors obtained using the item recall measure from the number of errors obtained using the strict measure).[4] This yields a measure of order accuracy that is independent of an individual's overall recall ability. When scored using the order accuracy measure, the detrimental PSE on order memory remained (Wickelgren, 1965d). Further, the term 'strict' was replaced with 'correct-in-position' to reflect the idea that the results observed when using this measure are influenced by the order constraints that the serial recall task places on recall performance.

As mentioned previously, Fallon et al. (1999; Experiment 1) found no differences in order memory for similar as compared to dissimilar lists when scored using the correct-in-position criterion (see Table 3.2). However, when scored using the order accuracy measure, the standard PSE was found, such that order memory was better for dissimilar as compared to the phonemically similar lists (refer Table 3.3). Further, Poirier and Saint-Aubin (1996) found that order memory was impaired for phonemically similar as compared to dissimilar lists of items when measured using the order accuracy criterion. However, when performance was measured for item recall, no differences in the recall of item information were observed (see Table 3.3). Therefore, consistent with previous research (e.g., Crowder, 1979; Watkins et al., 1974;

---

[4] Please note, the measure of order accuracy that is commonly used in current research (e.g., Fallon et al., 1999) is the number of items recalled in the correct position divided by the total number of items recalled. This can be contrasted with Wickelgren's (1965d) method of calculating order accuracy (i.e., subtracting the number of errors obtained using the item recall measure from the number of errors obtained using the strict measure).

Wickelgren, 1965d), Poirier and Saint-Aubin (1996) suggest that similarity influences order memory, rather than the recall of item information.

Table 3.3 Research Findings into the PSE when Performance is scored using the Item Recall and Order Accuracy Measures.

	Scoring Criterion	
Research	***Item Recall***	***Order Accuracy***
Coltheart (1993)	Detrimental	
Fallon et al. (1999; Experiment 1)	Detrimental	Detrimental
Fallon et al. (1999; Experiment 2)	Detrimental	Detrimental
Fallon et al. (1999; Experiment 1)	Facilitative	Detrimental
Gathercole et al. (1982)	Facilitative	
Fallon et al. (1999; Experiment 2)	No Difference	Detrimental
Poirier & Saint-Aubin (1996)	No Difference	Detrimental
Watkins et al. (1974)	No Difference	
Wickelgren (1965d)	No Difference	Detrimental

Table 3.3 lists all of the studies into the effect that similarity has on the recall of item information when the experimental stimuli are words. Hence, although some studies have found no difference in the recall of item information for dissimilar as compared to phonemically similar lists (e.g., Fallon et al., 1999; Experiment 2; Poirier & Saint-Aubin, 1996; Watkins et al., 1974), other studies have failed to replicate this finding (see Table 3.3). For instance, Gathercole et al. (1982; see also Fallon et al., 1999; Experiment 1) re-analysed their results and found a beneficial effect of similarity for the recall of item information (see Table 3.3). Further, other researchers have found the reverse effect with a decrease in the recall of item information for phonemically similar as compared to dissimilar lists. For example, Coltheart (1993; see also Fallon et al., 1999; Experiment 2) found that regardless of whether performance was scored for

item or order memory, the detrimental effect of similarity on STM performance remained. However, as Coltheart points out, most of the research into the effect that similarity has on the recall of item information have used word lists that were constructed from a limited stimulus pool (i.e., a closed word pool). Hence, the suggestion is that the contradictory findings observed for the recall of item information may be due to the *size* of the stimulus pools used to construct the lists.

3.5. Does the size of the word pool make a difference?

To examine whether the size of the stimulus pools used to construct the experimental lists influences the recall of item information, Coltheart (1993) constructed stimulus lists using either a closed or open word pool (i.e., words that are sampled without replacement from a large set). Coltheart found that regardless of the size of the stimulus pools used to construct the lists, similarity impaired the recall of item information (see Table 3.4).

Table 3.4 Research into the PSE when Performance was measured for Item Recall and the Stimuli were Chosen using either a Closed or Open Word Pool.

Research	**Word Pool**	**Item Recall**
Coltheart (1993)	Closed	Detrimental
Fallon et al. (1999; Experiment 2)	Closed	Detrimental
Fallon et al. (1999; Experiment 2)	Closed	No Difference
Coltheart (1993)	Open	Detrimental
Fallon et al. (1999; Experiment 1)	Open	Detrimental
Gathercole et al. (1982)	Open	Facilitative
Fallon et al. (1999; Experiment 1)	Open	Facilitative
Poirier & Saint-Aubin (1996)	Open	No Difference
Watkins et al. (1974)	Open	No Difference

However, it should be noted that even when the stimuli are chosen using an open word pool, the findings across studies for the recall of item information are still contradictory. For instance, in comparison to Colheart's (1993) finding of a detrimental effect of similarity on the recall of item information, Watkins et al. (1974) found no differences between phonemically similar and dissimilar lists of words (see Table 3.4). In contrast, Gathercole et al. (1982) found a facilitative effect of similarity on the recall of item information (see Table 3.4). Hence, the contradictory findings in the literature into the effect that similarity has on the recall of item information cannot solely be attributed to the size of the stimulus pools used to construct the lists. Fallon et al. (1999), proposed an alternative suggestion to account for these contradictory research findings. According to Fallon et al. the differential results reported in the STM literature are related to how similarity has been operationally defined.

3.6. Operationally defining similarity and its effect on the recall of item information

Fallon et al. (1999) examined whether operationally defining similarity in different ways influenced the recall of item information. They found that similarity impaired the recall of item information when the stimulus lists consisted of words that overlapped on a large number of phonemes, but did not rhyme (e.g., *ham, mass, map* and *had*). This detrimental effect of similarity on item recall remained, regardless of whether the stimulus lists were constructed using a closed or open word pool (refer Table 3.5). In comparison to dissimilar lists, Coltheart (1993) also found a detrimental effect of similarity for item recall with stimulus lists that overlapped on both the rhyme and other units (e.g., *cat, rat, cab* and *rag*). However, contrary to Coltheart's (1993) results, Watkins et al. (1974) found no item recall differences between dissimilar and similar lists when the phonemically similar lists were constructed using words that overlapped on both the rhyme and other units: although Watkins et al. used seven-item lists, whereas Coltheart used five-item lists per trial. Hence, the null finding observed by Watkins et al. may have been a result of task difficulty, rather than a true memorial effect. For instance, when order memory was scored using the correct-in-position measure, Watkins et al. found mean proportions correct of between *.31* and *.37* for the similar and between *.22* and *.37* for the dissimilar lists[5]. In comparison Coltheart (1993) found mean proportions correct of between *.62* and *.69* for the similar and between *.83* and *.95* for the dissimilar lists. Therefore, it is reasonable to suggest that regardless of whether the stimulus lists share some rhyme units in common, or share a high number of overlapping phonemes but do not rhyme, similarity impairs the recall of item information.

Table 3.5 Research into the PSE when Performance was measured for Item Recall. Depicting the Size of the Stimulus Pools Used to Select the Lists and How Similarity Has Been Operationally Defined.

Research	Stimuli Defined	Pool	Item Recall
Fallon et al. (1999; Experiment 2)	Overlap (None Rhyming)	Closed	Detrimental
Fallon et al. (1999; Experiment 1)	Overlap (None Rhyming)	Open	Detrimental
Coltheart (1993)	Overlap (Some Rhyming)	Open	Detrimental
Coltheart (1993)	Overlap (Some Rhyming)	Closed	Detrimental
Watkins et al. (1974)	Overlap (Some Rhyming)	Open	No Difference
Fallon et al. (1999; Experiment 2)	Rhyming	Closed	No Difference
Poirier & Saint-Aubin (1996)	Rhyming	Open	No Difference
Gathercole et al. (1982)	Rhyming	Open	Facilitative
Fallon et al. (1999; Experiment 1)	Rhyming	Open	Facilitative

In contrast, Fallon et al. (1999; Experiment 1; see also Gathercole et al., 1982) found an item recall advantage for stimulus lists that rhymed (e.g., *bog, hog, dog* and *log*) and were chosen using an open word pool (see Table 3.5). Nevertheless, Poirier and Saint-Aubin (1996) found no difference in the recall of item information between

[5] Please note, the Watkins et al. (1974) article did not include the means for the recall of item information. Thus, the means when performance was scored using the correct-in-position measure have been quoted.

dissimilar and rhyming lists of stimuli (see Table 3.5). However, except for a single one-syllable word, Poirier and Saint-Aubin used polysyllabic French words to construct their stimulus lists. In comparison, Fallon et al. (1999; see also Gathercole et al., 1982) constructed their stimulus lists with one-syllable English words. As such, it is difficult, if not impossible, to compare research findings across studies that have used one-syllable English words, with a study that has used polysyllabic French words and conclude that the results are contradictory. Therefore, the best explanation that can be generated for the inconsistencies that have been found in the research literature for the effect that similarity has on the recall of item information is that these discrepancies are a result of how similarity has been operationally defined (see Table 3.5). For instance, when the stimulus lists share a high number of overlapping phonemes but do not rhyme or share some rhyme units in common, a detrimental effect of similarity on the recall of item information is reported. However, when the stimuli share an English rhyme ending and are chosen using an open word pool, similarity facilitates the recall of item information (Fallon et al., 1999; Gathercole et al., 1982).

When the stimulus lists are constructed using a closed word pool, the beneficial effect that rhyming lists have on item recall is absent (see Table 3.5). For instance, Fallon et al. (1999; Experiment 2) found that when the stimulus lists were constructed using a closed word pool, item recall was similar for dissimilar as compared to rhyming lists. Nairne and Kelley (1999; p.46) suggest that when the stimuli are repeatedly sampled from a small set (closed word pool), "…one increases the probability that cross-trial confusions will occur (i.e., proactive interference)."

One way of assessing whether proactive interference is influencing performance when the stimulus lists are chosen using a closed word pool, is to compare the proportion of order errors observed for these lists with those observed when the stimulus lists are chosen using an open word pool. For example, in the Fallon et al. (1999) study when the stimulus lists were constructed using a closed word pool, order memory was *.90* for the dissimilar and *.71* when the stimulus lists rhymed. An identical pattern was observed when the stimulus lists were constructed using an open word pool (i.e., *.90* and *.71* respectively). In effect, when the stimuli are chosen using a closed word pool, it is impossible to tell whether the recall errors are order errors or cross-trial confusion errors. Hence, both of these types of errors would be classified as order errors. However, when the stimuli are chosen using an open word pool, these cross-trial confusion errors are classified as such. Therefore, using an open word pool to construct the stimulus lists yields a purer measure of the proportion of order errors observed. Hence, if cross-trial confusion errors are influencing the number of items recalled, then a larger number of order errors should be observed when the stimuli are chosen using a closed as compared to open word pool. Therefore, the findings observed in the Fallon et al. (1999) study suggest that the differences in item recall levels observed in the research literature when the stimulus lists are chosen using a closed as compared to open word pool cannot be attributed to an increase in cross-trial confusion errors.

A further two explanations have been proposed to account for the beneficial effect that choosing stimuli from a closed, as compared to open word pool, has on the recall of item information. For instance, Roodenrys and Quinlan (2000) suggest that when the stimuli are chosen using a closed word pool, the redintegration process is restricted to a small set of items. Hence, there is less competition at output from long-term representations that are similar to the target item as compared to when the stimuli are chosen using an open word pool. Alternatively, it may be the case that speech production plays a greater role in the recall of list items when the stimuli are chosen using an open word pool. For instance, if it is the case that speech production influences recall performance, then overall item recall should be higher when the stimulus lists are chosen using a closed word pool as compared to when the stimuli are 'new' on every trial. Comparisons between studies that have measured the recall of item information for dissimilar lists suggest that item recall is higher when the stimuli are chosen using a closed (Coltheart, 1993; .96; Fallon et al., 1999; .86), as compared to open word pool (Coltheart, 1993, .85; Fallon et al., 1999; .65). As such, these findings lend support to the idea that when the task is serial recall and the stimulus lists are chosen using an open word pool, the recall of item information may be influenced by speech output constraints.

3.7. Summary

A considerable amount of research has been conducted on the influence that similarity has on STM performance. In summary, earlier research suggested that similarity influenced the order in which items were recalled, rather than the number of items recalled (Murdock, 1976; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d). However, recent research has questioned this view. For instance, although the detrimental effect that similarity has on order memory is a robust finding in the research literature, when performance is measured for item recall, the results are contradictory. Fallon et al. (1999), suggest that the effect that similarity has on the recall of item information depends on how phonological similarity has been operationally defined. For example, when the stimulus lists share some rhyme units in common (e.g., Coltheart, 1993) or share a high number of overlapping phonemes but do not rhyme (e.g., Fallon et al., 1999) a detrimental effect of similarity on the recall of item information is observed. However, when the stimuli share an English rhyme ending and are chosen using an open word pool, similarity facilitates the recall of item information (Fallon et al., 1999; Gathercole et al., 1982). Based on this finding, Fallon et al. (1999) suggest that similarity differentially influences item and order memory, in that although similarity can have a facilitative effect on the recall of item information, it has a detrimental effect on order memory.

An alternative suggestion proposed to account for the beneficial effect that sharing a rhyme ending has on the recall of item information is that this item recall advantage is due to phonemic overlap (Fallon et al., 1999). For instance, when the stimuli rhyme, each word in a list shares two phonemes with every other list item. However, when words do not rhyme, each word may share two phonemes with some items, but only one phoneme with others. Hence, rhyming lists share a greater number of overlapping phonemes than do phonemically similar non-rhyming stimulus lists. The current work will investigate the idea that the beneficial effect of similarity on the recall of item information may be due to phonemic overlap. Further, an under researched area in the STM literature is the effect that phonological similarity has on STM when the stimuli are nonwords. Hence, the current thesis will also examine whether the effect that similarity has on both item and order memory for words persists when the experimental stimuli are nonwords.

3.8. Lexicality and the phonological similarity effect

The *lexicality effect* is the finding that STM performance is better for words as compared to nonwords or unfamiliar words (Hulme et al., 1991; Roodenrys et al., 1993). *Redintegration* is a term that has been used to describe the memory process that occurs prior to output. According to STM models that have incorporated a generic redintegration process, phonological representations that are stored in LTM can be used to fill in or redintegrate imperfect phonological traces held in STM (Brown & Hulme, 1995; Schweickert, 1993). Hence, STM performance should be better for words as compared to nonwords because there are no stored representations available to assist in the reconstruction of a partial trace when the stimuli are nonwords (Hulme et al., 1991; Hulme et al., 1997).

Recently, researchers interested in investigating the influence that similarity has on STM performance have turned their attention to nonwords (Fallon et al., in press; Gathercole et al., 2001; Lian et al., 2001). There are a number of theoretical reasons why investigations into the PSE using nonword stimuli are important. For instance, Saint-Aubin and Poirier (2000, p. 333; see also Brown et al., 2000; Burgess & Hitch, 1992) suggest that, "…given that no adequate long-term representations are available for nonwords, the reconstruction process, for all practical purposes, is thought not to operate for these items". The idea that, unlike when the stimuli are nonwords, the redintegration process aids in the recall of words is used to explain why memory span is lower for nonwords. However, Fallon et al. (in press) have recently suggested that the redintegration process operates for both words and nonwords. Therefore, if the same processes are involved in word and nonword recall, as suggested by Fallon et al. (in press), then extant STM models would need to be modified to reflect this. Further, given that a number of researchers suggest that the PSE arises during the redintegration process (Saint-Aubin & Poirier, 2000), research findings that suggest that the PSE is also present when the experimental stimuli are nonwords may help illuminate the level at which changes to extant STM models are necessary.

3.8.1. Current research findings on the phonological similarity effect using the serial recall task with nonword stimuli

Currently there are only a handful of studies that have examined the effect that similarity has on STM with nonwords. In one study, Drewnowski (1980; Experiment 1) created four lists of stimuli that consisted of consonant-vowel (CV) nonwords. In the similar conditions, one set of nonwords (i.e., consonant only condition), shared a common vowel and different initial consonants (e.g., *gah, sah,* and *fah*), the other set (i.e., vowel only condition), shared a common initial consonant and different vowels (e.g., *dih, dah,* and *diy*). In the redundant conditions, all of the nonwords presented in a given list did not share any common phonemes. What makes these conditions interesting is that, even though all of the list items were phonemically dissimilar, in one condition (i.e., the redundant vowel, consonant only condition), the position of the vowels did not vary across trials. For example, if the first trial presented to participants consisted of the nonwords, g*ah*, s*oy*, f*eh*, the next trial might have consisted of the items, f*ah*, d*oy*, z*eh*. In the other redundant condition (i.e., redundant consonant, vowel only condition), the position of the consonants did not vary across trials (e.g., *b*ih, *f*ah, *d*iy, as compared to *b*iy, *f*ih, *d*eh). Participants were told the order in which the redundant vowels (consonant only condition) and the redundant consonants (vowel only condition) would occur, *before* the first experimental lists were presented. Thus, as Drewnowski (1980; p. 179) suggests, in the redundant vowel, consonant only condition "….vowel sounds were effectively prevented from contributing to string recall because the same sequence of vowels was repeated from trial to trial" and vice versa for the redundant consonant, vowel only condition.

Participants were visually presented with six item lists and their task was to write down the nonwords in the order in which they were presented. Thus, strict serial recall instructions were employed. For the similar conditions, Drewnowski (1980) found that, order memory was better for nonwords that consisted of a different vowel (i.e., vowel as compared to the consonant only condition). Further, when performance was compared across the redundant conditions, the same pattern of results was observed (i.e., order memory was better for the redundant consonant, vowel only as compared to the redundant vowel, consonant only condition). This is an important finding in that all of the vowels and consonants in redundant conditions were phonemically distinct. Hence, "the mere presence of acoustically distinct vowels in the stimulus string is … not sufficient for the improvement of syllable recall" (Drewnowski, 1980; p.182). Further, when performance was compared across the similar and redundant (i.e., dissimilar) conditions, no differences in correct-in-position recall were observed. However, it should be noted that in the redundant conditions, the same consonant or vowel was presented in the same position across each list, whereas in the similar conditions, the positions in which the phonemes within items were presented differed across trials. Thus, the results observed in the redundant conditions may have been influenced by what is commonly known as the Hebb repetition effect. This is the finding that when stimuli are repeatedly presented in the same position in a list, recall accuracy increases as the number of trials increase (Hebb, 1961).

Besner and Davelaar (1982) also investigated the influence that similarity has on STM with nonwords. Using a closed pool (i.e., ten nonwords in each set), Besner and Davelaar constructed four-item lists that either rhymed or were phonemically dissimilar. They also constructed lists of confusable as compared to non-confusable pseudohomophones. Pseudohomophones are letter strings that are nonwords when presented visually, but sound like real words when pronounced (e.g., *phood, brued* and *chood*). Using the correct-in-position criterion, Besner and Davelaar found that order memory was higher for dissimilar as compared to rhyming lists of nonwords (see Table 3.6). Further, order memory was also higher for phonemically dissimilar pseudohomophones in comparison to pseudohomophones that rhymed. This finding of a detrimental effect of similarity on order memory is consistent with previous research that has employed letters and words in comparison to nonword lists (e.g., Conrad & Hull, 1964; Baddeley, 1966).

A more recent study using nonword stimuli was conducted by Lian et al. (2001). Lian et al. (2001; Experiment 1A) manipulated the lexicality of the stimuli (i.e., words vs. nonwords), the similarity of the list items (i.e., either dissimilar or lists that shared the vowel), and associative value (i.e., items that were rated as being either more or less wordlike in terms of reaction time). Consistent with research into the effect that lexicality has on STM (e.g., Hulme, Roodenrys, Brown & Mercer, 1995), Lian et al. (2001) found that memory span was lower when the experimental stimuli were nonwords as compared to lists of words. Further, Lian et al. (2001; see also Besner &

Davelaar, 1982), found a detrimental effect of similarity on order memory, such that phonemically similar lists were recalled in the correct order less often than dissimilar lists of nonwords (see Table 3.6), but only when the stimulus lists were said to be high in associative value (i.e., rated as more wordlike). When the nonwords were rated as being low in associative value, the reverse effect was found with better performance for similar as compared to distinct lists of nonwords.

Table 3.6 Experimental Findings on the PSE with Methodological Differences in Stimuli Presentation and Recall Techniques when the Experimental Stimuli were Nonwords.

	Presentation	Recall	Scoring
Research			***Correct-in-position***
Drewnowski (1980)	Visual	Written	No Difference
Besner & Davelaar (1982)	Visual	Verbal	Detrimental
Lian et al. (2001; Experiment 1A)			
High Associative Value	Auditory	Verbal	Detrimental
Low Associative Value	Auditory	Verbal	Facilitative
Lian et al. (2001; Experiment 1B)			
High Associative Value	Auditory	Verbal	Detrimental
Low Associative Value	Auditory	Verbal	No Difference
Gathercole et al. (2001)	Auditory	Verbal	Detrimental
Fallon et al. (in press)	Visual	Written	Facilitative

In an attempt to replicate the above mentioned findings, Lian et al. (2001; Experiment 1B) conducted a further experiment in which similarity was defined in terms of a shared rhyme unit. The results replicated their earlier study when the nonwords were rated high in associative value (i.e., Experiment 1A), such that memory span was higher for dissimilar as compared to the phonemically similar lists of nonwords. However, no effect of similarity on order memory was found for nonwords that were rated as being low in associative value. The explanation given for these contradictory results was that nonwords that are classified as being high in associative value are more similar to real words than those that are rated as low in associative value. Accordingly then, the more wordlike an item is, the easier an LTM representation for an item can be activated. This idea is consistent with previous research that suggests that the more wordlike a nonword is rated, the more accurately it is recalled (e.g., Gathercole, 1995; Gathercole, Willis, Emslie & Baddeley, 1991; Gathercole & Martin, 1996; Metsala, 1999).

Recently, Gathercole et al. (2001) used an open word pool to construct lists of words and nonwords that were either dissimilar (i.e., syllables with a different vowel and consonants that rarely repeated across a list), or phonemically similar (i.e., syllables with the same vowel and different consonants across all list items). Consistent with Lian et al.'s (2001) research, Gathercole et al. (2001; Experiments 3A and 4A) found that memory span was lower when the stimulus lists consisted of nonwords in comparison to words. Further, when scored using the correct-in-position measure, order memory was higher for dissimilar as compared to the phonemically similar lists of items (refer Table 3.6).

Finally, Fallon et al. (in press) recently looked at the differences between word and nonword recall with stimulus lists that were either phonemically similar (i.e., rhyming) or dissimilar (refer Table 3.6). Contrary to Gathercole et al.'s (2001) results, Fallon et al. found that when scored using the correct-in-position measure, performance was higher for similar as compared to dissimilar lists of nonwords. This finding of a facilitative effect of similarity for rhyming nonwords persisted, regardless of whether participants were silent or performed articulatory suppression during the recall phase (Fallon et al., in press; Experiments 1 & 2).

Given the suggestion that similarity differentially influences order memory and the recall of item information (i.e., Fallon et al., 1999), and the idea that the correct-in-position criterion is not a pure measure of order memory (Wickelgren, 1965d), Fallon et al. (in press) scored performance using both the item recall and order accuracy measures. Consistent with previous studies that have used an open word pool to construct rhyming lists (e.g., Fallon et al., 1999; Experiment 1; Gathercole et al., 1982), Fallon et al. (in press) found a facilitative effect of similarity for the recall of item information. Further, when performance was scored using the order accuracy measure, order memory was better for dissimilar as compared to rhyming lists of nonwords. Therefore, consistent with studies that have used words, similarity influences a participant's ability to correctly order presented list items (e.g., Murdock, 1976; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d). Further, these findings also suggest that the phonological similarity of nonwords as with words, can, under certain conditions, have a beneficial effect on the recall of item information (e.g., Fallon et al. 1999; Fallon et al. in press; Gathercole et al., 1982).

3.8.2. Summary

To date, research on the influence that similarity has on STM performance with nonwords as compared to words is limited. In summary, consistent with previous word studies (e.g., Baddeley, 1966; Coltheart, 1993; Li et al., 2000; Watkins et al., 1974), similarity has been found to impair order memory, such that dissimilar lists are recalled in the correct order more often than phonemically similar lists of nonwords (Besner & Davelaar, 1982; Gathercole et al., 2001). However, a number of results are inconsistent with the abovementioned findings. First of all, in comparison to the dissimilar lists, Lian et al. (2001) found a facilitative effect of similarity for nonwords that were rated as being low in associative value (Experiment 1A; see also Fallon et al., in press). Second, Lian et al. (2001; Experiment 1B; refer Table 3.6) found no differences in order memory, regardless of whether nonwords that were rated as being low in associative value, were phonemically dissimilar or shared a rhyme ending. However, it should be noted that the stimuli used in the Lian et al. (2001) study were based on Norwegian language constraints, whereas other nonword studies have constructed nonword lists that are based on English language-constraints.

Further, one suggestion proposed by Lian et al. (2001) to account for these inconsistent results is that nonwords that are rated as being low in associative value (i.e., less wordlike) impose a higher memory load when recall is dependent on phoneme representations. According to this view, those items that are phonemically dissimilar should be harder to recall than similar lists of nonwords. Hence, the finding that memory span was higher for phonemically similar as compared to dissimilar lists of nonwords (Lian et al., 2001; Experiment 1A). However, it is unclear how this memory load argument can be used to account for the finding that when the nonwords rhymed (Experiment 1B) as compared to the phonemically dissimilar lists, no differences in STM performance were observed. Based on the memory load argument, one would expect performance for rhyming lists to be better than the phonemically dissimilar lists of nonwords: a prediction that is clearly inconsistent with the results observed in the Lian et al. (2001; Experiment 1B) study. Further, Lian et al. (2001) did not measure performance for item recall or use the order accuracy criterion to measure order memory. This is important in that performance measured using the correct-in-position criterion is not a pure measure of order memory as it does not take into account individual differences in item recall ability across conditions (Wickelgren, 1965d).

Currently, Fallon et al. (in press) have conducted the only study into the effect that similarity has on both the recall of item information (i.e., item recall measure of performance) and order memory (i.e., using the order accuracy measure of performance), when the experimental stimuli are nonwords. They found that when performance was measured using the order accuracy criterion, order memory was better for phonemically dissimilar as compared to rhyming lists of nonwords. This finding is consistent with studies that have used words as compared to lists of nonwords (e.g., Fallon et al., 1999; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d). The findings using the item recall measure of performance were also consistent with previous research with word lists (e.g., Fallon et al., 1999; Gathercole et al., 1982), in that similarity was found to facilitate the recall of item information.

In summary, the findings from the Fallon et al. (in press) study suggest that the effect that similarity has on both item and order memory is similar, regardless of whether the experimental stimuli are words or nonwords. Furthermore, most of these nonword studies have not examined the effect that similarity has on order performance with a measure of order memory that is not influenced by item recall ability. As such, more research needs to be carried out in this area before conclusions can be drawn as to whether the same (Fallon et al., in press) or different (Lian et al., 2001; Saint-Aubin & Poirier, 2000) processes are involved when recalling words as compared to nonwords.

3.9. Why look at the phonological similarity effect with nonwords?

There is one broad question that research into the effect that similarity has on STM with nonwords attempts to address, 'Are words and nonwords processed in the same way?' In other words, 'Are the same mechanisms involved in both word and nonword recall?' Currently a great debate rages as to whether words and nonwords are processed in the same, or a different way. For instance, according to Fallon et al. (in press) there is no reason to suggest that words and nonwords are processed differently. The idea that words and nonwords use the same recall processes is based on research findings that suggest that the redintegration process operates for both word and nonword recall (for an alternative view see Saint-Aubin & Poirier, 2000). In contrast, Lian et al. (2001; p. 289) suggest that "…when associative value is taken into account, we can conclude that words and nonwords are processed differently in short-term memory."

There are a number of reasons why care should be exercised if one is to conclude that words and nonwords *are* processed differently. Firstly, the current view of the PL is that it makes word learning and thus vocabulary acquisition possible (Baddeley et al., 1998). Also, as Gupta and MacWhinney (1997) suggest, for an individual who is learning a language, 'today's words were yesterday's nonwords'. Therefore, the idea is that the difference between word and nonword recall is one of degree. This places word and nonword recall on opposite ends of the same spectrum, with item familiarity as the variable influence on STM performance. The idea that word and nonword recall is influenced by the degree to which items are familiar is consistent with previous research findings. For instance, Hulme et al. (1991) found that memory span was higher for words as compared to nonwords with an English sound (e.g., *maffow*) or unfamiliar Italian words (e.g., *lago*). Further, Hulme et al. (1995) have also found that making nonwords more familiar to participants (i.e., using a pronunciation familiarisation task), improves serial recall performance. Other research which strengthens this claim is the finding that, the more wordlike a nonword is rated (Gathercole et al., 1991b; Gathercole, 1995; Gathercole & Martin, 1996; Metsala, 1999), or the quicker participants are able to associate a real word with a nonword (Lian et al., 2001), the more accurately the nonword is recalled.

A further issue is the influence that speech production processes have on STM performance. For instance, research to date has failed to separate phonological processing which includes phonological input and speech output (Hulme & Snowling, 1992; Snowling, Goulandris, Bowlby & Howell, 1986) from "…the phonological memory component (i.e., maintaining the phonological representation in the phonological loop)" (Bowey, 1997; p.298). This is important in that the influence that overt speech production processes have on STM performance when participants are required to verbally recall lists of stimuli has been proposed to account for the differences observed in the research literature when the stimulus lists are constructed using an open as compared to closed word pool (see Sections 3.5 and 3.6). In a recent study, Baddeley, Chincotta, Stafford and Turk (2002) separated the effect that speech production has on STM performance. Using both the serial recall and serial recognition tasks, they examined the effect that word length has on STM performance and found that although the word length effect on STM performance remained, "…the effect is substantially larger when tested by recall" (Baddeley et al., 2002; p.366). Further, recent research suggests that the ease of the articulatory transitions between experimental list items also influences STM performance (Murray & Jones, 2002; see also Service & Maury, 2003). Complimentary to this idea is the suggestion that when the stimuli are nonwords, STM may be further impaired by the influence that overt speech production has on STM performance (Gathercole et al., 1999). Hence, it is important to critically examine research into the effect that similarity has on STM performance with a task that is not constrained by the demands that overt speech production place on STM performance.

3.10. Investigating the phonological similarity effect without the influence that overt speech production has on STM performance

Although not a new proposal, research directed at understanding the mechanisms that are involved when performing a serial recall task has lead to the suggestion that these tasks are influenced by an individual's articulatory ability (Gathercole et al., 1999; Snowling & Hulme, 1989; Wells, 1995). For instance, Gathercole and Baddeley (1996) argue that when the task is nonword repetition (i.e., repeating a multisyllabic nonsense word), it is crucial that participant's have good articulatory skills if they are to perform well on such a task. This is due to the scoring criterion in that if a participant pronounces a single phoneme incorrectly, the entire utterance is scored as an error. Thus, the suggestion is that the results observed from STM tasks that do not require participants to overtly recall list items, are unaffected by factors such as speech production errors that may influence the results obtained, and as such the types of tasks may provide researchers with a purer STM measure (Gathercole et al., 2001).

Two tasks that do not require the overt recall of presented list items have been designed to examine the effect that similarity has on STM. However, each task has been designed to address specific issues within the literature. For instance, the 'free reconstruction of order' task has been used to investigate the effect that phonological similarity has on order memory when item information is made redundant (Neath, 1997). However, as Neath (1997; p. 262) admits,

> *"…even if one wanted to argue that the long-term free reconstruction of order task is nominally a pure measure of order memory because it does not, of itself, require the subject to remember item information, …nonetheless…subjects do remember and do use item information to complete the test. Because subjects use item information, the test is functionally not a pure measure of order memory."*

Nevertheless, the free reconstruction of order task is typically used as an example of a test that is free from the influence that the retention of item information has on order memory (Whiteman, Nairne & Serra, 1994). The second task that has been used to examine the effect that similarity has on STM performance when overt speech production is not required is called the *serial recognition task*. The utility of using tasks that do not require participants to verbalise presented list items has been demonstrated in neuropsychological studies with individual's that have speech production deficits, yet show intact STM performance (e.g., Martin & Breedin, 1992; Martin et al., 1999).

3.10.1. Separating STM for item information from memory for an item's position in a list

"Accurately reproducing the order of a list of items or set of events requires remembering two different kinds of information: information about the identity of each item, and information about the presentation order" (Neath, 1997; p. 256). In an attempt to separate item and order memory, Neath (1997) used what has been called the *free reconstruction of order task*. In this task, participants are presented with a list of items. At recall participants are re-presented with the same list of items, either in an alphabetical or a new random order. Hence, item information is available throughout the reconstruction process to aid in the reconstruction of order (Neath, 1997).

Nairne and Neumann (1993) conducted a study using the free reconstruction of order task to examine the effect that phonological similarity has on order memory. They presented participants with five-word lists that were either phonemically similar (i.e., shared the vowel) or phonemically dissimilar. A participant's task was to rate the presented words on a scale from 1 (*unpleasant*) to 3 (*pleasant*). Once all of the to-be-reconstructed lists had been presented, participants were given a 10-min distractor task. Participants were then presented with the original list items that had been printed on a new piece of paper in a random order. A participant's task was to place the words back in the order in which they had been originally presented. Nairne and Neumann found a beneficial effect of phonemic similarity for order reconstruction, in that order reconstruction was better for phonemically similar as compared to dissimilar lists of items. Based on these findings, Nairne and Neumann suggest that when lists can be easily discriminated from each other, such is the case when the items in one list share a unique feature, phonemic similarity should facilitate order memory. However, in this experiment, participants performed the free reconstruction of order task after a 10-min distractor task. Also, participants were required to reconstruct the order of multiple lists that were presented simultaneously at recall. Hence, caution should be exercised when comparing the results observed in the Nairne and Neumann (1993) study with experiments that require participants to firstly, perform a task without delay, and secondly, recall items after each individual list has been presented.

One study that examined order memory with the order reconstruction task immediately after each list had been presented, and at shorter retention intervals, was conducted by Nairne and Kelley (1999). They designed the study primarily to explore the idea that with the right experimental manipulations, phonemic similarity can facilitate order memory. Nairne and Kelley varied the retention interval (i.e., 2, 8, or 24 second intervals), whether similarity was blocked across experimental sessions (i.e., either presenting participants with phonemically similar or dissimilar lists first, and vice versa, or presenting the lists randomly throughout the experimental session) and the size of the stimulus pools used to select the list items (i.e., open as compared to closed word pool). They found that regardless of the size of the stimulus pool or whether the experimental conditions were blocked, phonemic similarity had a detrimental effect on order memory at a 2 sec interval. However, no effect of phonemic similarity on order memory was found when the interval was increased to 8 secs.

When memory for order was measured after a 24 sec interval, a different pattern of results emerged depending on whether an open or closed word pool was used to select the stimulus lists (Nairne & Kelley, 1999). For instance, when the stimuli were chosen using an open word pool (Experiments 1 & 2), phonemic similarity had a beneficial effect on order memory. However, similarity impaired order memory when the stimuli were chosen using a closed word pool (Experiment 3). Nairne and Kelley explain this reversal of the PSE in terms of the discriminability of items along the list dimension. For instance, Nairne and Kelley argue that the within-list dimension drives performance when testing participants immediately after list presentation, such that the standard PSE on order memory should be observed. However, at longer delays, memory is primarily driven by an individual's ability to discriminate items on the list-dimension. Hence, when items are novel on every trial (i.e., open word pool), individuals are able to use a common feature (e.g., a rhyme ending or common vowel) to aid in locating the correct list to-be-reconstructed. In contrast, when the stimulus lists are constructed using a closed word pool, participants can no longer use the list dimension as a cue in which to aid reconstruction performance. Hence, the standard PSE on order memory should be observed.

However, care needs to be taken when comparing performance across studies that have used the free reconstruction of order task as compared to the more traditional serial recall task. First of all, when performing a free reconstruction of order task, participants are presented with the items *during* the recall phase. This is not the case when the task is serial recall where participants are required to recall both item and order information. Also, when performing the free reconstruction of order task, participants are allowed to reconstruct the lists in any order they choose. In contrast, the serial recall task requires that participants recall the first item presented first, and each subsequent item sequentially. Finally, as compared to the traditional serial recall task, given that item information is presented to participants *during* the recall phase, the results when using the free reconstruction of order task may be influenced to a greater extent by guessing (Healy, 1974). An alternative task that does not require participants to verbally recall presented list items, yet is more similar to the standard serial recall task, is the serial recognition task.

3.10.2. The serial recognition task

The serial recognition task was specifically designed to assess an individual's memory for the position in which stimuli are presented, independent of an individual's overt speech production ability (Campbell & Butterworth, 1985; Martin & Breedin, 1992; Martin et al., 1999). This task involves presenting a list of items and then re-presenting either the same items in the same order or the same item's in a different order (i.e., usually two adjacent items are swapped). A participant's task is to say whether the items re-presented were in the 'same' (e.g., *log, bog, hog* and then *log, bog, hog*) or a 'different' (e.g., *log, bog, hog* and then *bog, log, hog*) order. Because this task is resistant to speech production errors, Gathercole et al. (2001) have recently suggested that in comparison to the standard serial recall task, the serial recognition task may be a better measure of STM. This is because unlike the standard serial recall task, serial recognition does not require participants to overtly articulate the presented items.

Two recent studies have used the serial recognition task to look at the influence that similarity has on order memory. In one study, Gathercole et al. (2001) presented participants with lists of one-syllable words, followed by a one second pause. Participants were then presented with the same items in the same order or with a list in which two adjacent items had been transposed. They found that phonemically dissimilar items were recognised as being either in the 'same' or a 'different' order more accurately than when the stimulus lists were phonemically similar, regardless of whether the stimuli were words or nonwords. Further, Lian et al. (2001) also found that recognition performance was higher when the stimuli were phonemically dissimilar as compared to similar lists of words. These findings are consistent with previous research into the effect that phonological similarity has on order memory when performance is scored using the order accuracy criterion (e.g., Fallon et al., in press; Wickelgren, 1965d).

In contrast to the findings from the Gathercole et al. (2001) study, using the serial recognition task, Lian et al. (2001) found no differences between dissimilar and phonemically similar lists of nonwords that were rated as being low in associative value. However, the nonwords Lian et al. used in this study were drawn from a previous study that had found a facilitative effect of similarity for order memory when the task was serial recall (see section 3.8.2 for further criticisms regarding the design of this study). Thus, this finding of no difference in order memory for phonemically similar as compared to dissimilar lists of nonwords is not entirely convincing.

3.10.3. Summary of research findings into the phonological similarity effect with tasks that do not require the overt production of presented list items

Gathercole et al. (1999) have recently suggested that the results observed when using a serial recall task to measure STM performance may be influenced by an individual's articulatory ability. As such, the free reconstruction of order and serial recognition tasks have been used to examine the effect that similarity has on order memory, independent of the influence that overt speech production may have on the research findings. Using the free reconstruction of order task, Nairne and Neumann (1993; see also Nairne & Kelley, 1999) have found a facilitative effect of similarity on order reconstruction. Nairne and Kelley (1999) suggest that when the items in a list can be easily discriminated from the items that were presented in a different list, phonemic similarity can have a facilitative effect on order memory. For instance, similarity has been found to facilitate order memory when lists are drawn from unique categories, such as different birds in one list and flowers in a subsequent list (Nairne, 1990b), and when the stimulus lists share a rhyme ending (Fallon et al., 1999). However, even when the stimuli in a list can be classified as belonging to a particular category (e.g., shared rhyme ending), similarity will only facilitate item and order memory when the category *uniquely* identifies an item as belonging to a particular list (Nairne & Kelley, 1999). Hence, the current view is that the effect that similarity has on item and order memory is dependent on how similarity is operationally defined, both within a list and between the lists used in any one experimental session (Fallon et al., 1999).

One way of experimentally manipulating the list dimension is to use either a closed or open word pool to select the stimulus lists (Fallon et al., 1999; Nairne & Kelley, 1999). An interesting quandary emerges when experimenters are choosing the size of the stimulus pools that will be used across experiments. For instance, Nairne and Kelley (1999) argue that repeatedly sampling items from a small set (i.e., closed word pool) increases the chances that an individual's performance will be contaminated by cross trial confusion errors. In contrast, when the stimuli are not repeated in an experimental session (i.e., open word pool), an individual's performance may be influenced by their articulatory ability (Snowling & Hulme, 1989; Wells, 1995). At present, the extent to which articulatory ability influences item and order memory for phonemically similar as compared to distinct lists of items has *not* been explored in detail.

Two recent studies have examined the effect that similarity has on STM performance with a task that is not influenced by overt speech production processes. With the exception of one of the findings from the Lian et al. (2001) study that used lists of nonwords that were rated as being low in associative value (see section 3.10.2.), the findings suggest that when the influence that speech production has on STM performance is controlled, the standard detrimental effect that similarity has on order memory is observed. In other words, order memory appears to be better for phonemically dissimilar as compared to similar lists of items, regardless of whether the experimental stimuli are words or nonwords (e.g., Gathercole et al., 2001).

3.11. Outline of the experimental chapters

The experimental chapters in the current thesis are in manuscript format. Each chapter is self-contained in that the experiments were designed to investigate a unique aspect with regards to the influence that similarity has on STM. There are four broad issues that the current thesis was designed to address. The first issue is whether the phonemic overlap argument can account for the item recall advantage that has been observed when list items rhyme (Study 1 - Chapter 4). Study two was designed to examine whether the effect that phonological similarity has on both the recall of item information and memory for an item's position in a list persists when the experimental stimuli are nonwords (Chapter 5). Regardless of whether the experimental stimuli were words or nonwords, study three (Chapter 6) was designed to examine the effect that phonemic similarity has on order memory, once the demands that overt speech production processes have on STM performance are removed. Finally, the explanations that current STM models generate for the PSE have been critically examined throughout each of the abovementioned studies.

PRELUDE TO CHAPTER 4

When the stimuli are words, early research suggested that similarity has its effect on STM performance at an order rather than item level (Murdock, 1976; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d). When scored using the correct-in-position measure, research findings are consistent in that phonemically dissimilar lists are recalled better than similar lists of items (e.g., Baddeley, 1966; Coltheart, 1993; Gathercole et al., 1982; Li et al., 2000; Poirier & Saint-Aubin, 1996; Watkins et al., 1974). However, when performance is scored using the item recall criterion, the results are ambiguous with some studies finding a facilitative effect (e.g., Fallon et al., 1999), some no effect (e.g., Poirier & Sain-Aubin, 1996), and still others, a detrimental effect (e.g., Coltheart, 1993) of similarity on the recall of item information. Currently there are two competing suggestions to account for these contradictory findings. For instance, Fallon et al. (1999) suggest that there is something special about the rhyme unit, in that the rhyme acts as a cue to facilitate item recall. An alternative suggestion proposed to account for the item recall advantage observed for rhyming lists of stimuli is that this facilitative effect is due to phonemic overlap (Fallon et al., 1999).

The initial sets of experiments were specifically designed to examine whether the rhyme unit can act as a cue to facilitate item recall, or whether this item recall advantage is due to phonemic overlap. This is an important issue in that non-linguistic STM models are based on the distinctiveness assumption (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a) which is the idea that it is the *distinctiveness* of a memory trace in relation to other presented list items that is important for recall (Gillund & Shiffrin, 1984; Luce, 1959; Nosofsky, 1986). In contrast, psycholinguistic models of STM (e.g., Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) assume that linguistic processes operate at the sub-syllabic level to influence STM performance. Hence, the current sets of experiments were also designed to assess the utility of STM models based on the explanations that each model generates for the observed experimental findings.

4. INVESTIGATING THE PHONOLOGICAL SIMILARITY EFFECT: SYLLABLE STRUCTRE AND THE POSITION OF COMMON PHONEMES*

By

Lisa M. Nimmo and Steven Roodenrys

University of Wollongong
Wollongong, Australia

Address for correspondence:
Lisa Nimmo, Department of Psychology, University of Wollongong
Northfields Ave, Wollongong, 2522, Australia
Phone: (+61) 242 214 454, Fax: (+61) 242 214 163
Email: lmn02@uow.edu.au

Running head: Phonological similarity, serial recall and STM models

* This paper is currently in press (Journal of Memory and Language)

4.1. Abstract

The aim of the present research was to determine whether the effect that phonological similarity has on immediate serial recall is influenced by the consistency and position of phonemes within words. In comparison to phonologically dissimilar lists, when the stimulus lists rhyme there is a facilitative effect on the recall of item information and a detrimental effect on order memory (Experiment 1). When stimuli share the initial consonant and vowel (Experiment 2) or the same initial and final consonant (Experiment 3), there is no beneficial effect of similarity for item information, coupled with a detrimental effect on order memory. Contrary to the predictions made by non-linguistic models of STM, the influence that similarity has on both the recall of item information and memory for the position of items in a list is dependent on which components of the items are shared within a list.

4.2. Introduction

One of the theoretically most influential findings in studies of verbal short-term memory (STM) is the phonological similarity effect (e.g., Baddeley, 1966; Conrad et al., 1966; Conrad & Hull, 1964; Gathercole et al., 2001; Laughery & Pinkus, 1966; Schweickert et al., 1990; Sperling & Speelman, 1970): The finding that immediate serial recall is worse if words sound similar to each other than if they do not.

It has been known for some time (e.g., Wickelgren, 1965d) that effects on order memory, rather than item memory mediate the phonological similarity effect. The same numbers of words are recalled in similar and dissimilar lists, but items in similar lists are more likely to be recalled in the wrong order (see also Poirier & Saint-Aubin, 1996; Watkins et al., 1974). This phonological similarity effect is so robust that any successful model of STM must be able to explain it.

Until recently little consideration has been given to what is meant by similarity and the possibility that different operational definitions may contribute to inconsistent findings in the literature. In some studies phonological similarity has been operationally defined as lists of rhyming words (e.g., Gathercole et al., 1982; Poirier & Saint-Aubin, 1996), while other studies have used lists of single syllable words with a common vowel and some overlap in the consonants (e.g., Coltheart, 1993; Watkins et al., 1974). Fallon et al. (1999) directly compared the recall of lists of rhyming words, a phonemically similar condition in which items shared common phonemes but did not rhyme, and a dissimilar condition. They found that although both the rhyming and the phonemically similar condition showed impaired order memory compared to a dissimilar condition, the recall of item information was actually enhanced in the rhyming condition (see also Gathercole et al., 1982). That is, more items were recalled, albeit in the wrong order, when all of the items in a list rhymed than when the words in a list were phonologically dissimilar.

It could be argued that a rhyming list of consonant-vowel-consonant (CVC) words is quantitatively more similar than a list of words with a common vowel plus some consonant overlap, because every word shares two phonemes (the _VC segment) with every other word in a list when they all rhyme. The finding that item recall is better in the rhyming condition (where there is more overlap) implies that the nature of the similarity effect may be determined by the amount of phonemic overlap. However, this raises the question of whether the position of the phonemic overlap within the words is important. Lists of CVC words where all items share two of the other phonemes (i.e., consonant-vowel, CV_, or consonant-consonant, C_C) can be regarded as having an equivalent degree of similarity as the rhyming (_VC) lists. However, there is evidence that the rhyme unit (_VC) is more tightly bound in STM as when a recall error occurs, it is the pairing most likely to remain intact when the stimuli are CVC nonwords (e.g., Treiman & Danis, 1988). This suggests that the VC may have a special role in STM and so the effects observed with rhyming lists may not be apparent with lists sharing the CV_ or the C_C. The aim of the following experiments was to evaluate if these conditions are functionally equivalent, and in doing so to evaluate the predictions of two different classes of STM models.

4.2.1. Models of short-term memory

Models of STM can be broken into two general classes. One class of models (e.g., Brown et al., 2000; Nairne, 1990a) views STM as a specialised memory mechanism and explains performance in terms of processes and principles that are applicable to almost any type of information. That is, they could model recall of objects, spatial locations, pictures or verbal material with only relatively minor modifications. These models stem from a tradition of memory research going back over 100 years and have been a prominent approach to understanding STM in the psychological literature.

An alternative, psycholinguistic perspective views performance on the serial recall task as being based on language processes (e.g., Gupta & MacWhinney, 1997; Martin et al., 1999). According to this view the processes underlying verbal serial recall are not simply general purpose mechanisms that exist to preserve a brief record of the immediate past, but instead reflect the operation of specialised language processing mechanisms whose primary purpose is to allow us to produce and comprehend spoken language. As such, performance may reflect constraints specific to verbal stimuli and use mechanisms that may be fundamentally different from those involved in remembering non-verbal material. What follows is a brief description of the two types of models, with some examples, and an explanation of the different predictions they make regarding the effect that phonemic similarity has on STM recall when it is defined as words sharing the CV_, the _VC or the C_C components.

4.2.1.1. Non-linguistic models of short-term memory

Many models of STM make few assumptions about the nature of the representations underlying serial recall performance. Items are often represented as a vector and the models are governed by general principles that pay no regard to any specific constraints related to the linguistic nature of the stimuli. As such, similarity in these models is a quantitative variable and without additional assumptions CVC items that share the _VC segment would be just as similar as items that share the CV_ or C_C.

A major concern of the non-linguistic models of STM has been to provide a mechanism to explain the retention of the order of list items, as this appears to be one of the major differences between short- and long-term memory. In many of these models the likelihood of recalling an item in the correct position is a function of its distinctiveness from all other items in the list. This reliance on distinctiveness explains the detrimental effect of similarity on memory for order in several models (Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a; Nairne & Kelley, 1999; Tehan & Fallon, 1999). For instance, according to the feature model (Nairne, 1988, 1990a), the effect that similarity has on order memory arises from the interpretation or "deblurring" of memory traces at retrieval. It is assumed that confusion arises when a degraded trace from primary (short-term) memory is compared with a set of traces from secondary (long-term) memory. Nairne (1990a) suggests that the selection of items from the 'secondary memory search set' is based on the ratio rule - the likelihood of recalling a presented item is a function of the relative similarity of the probe item to the similarity of all of the presented list items (Gillund & Shiffrin, 1984; Hintzman, 1986; Luce, 1959; Nosofsky, 1986). In most models, items are represented as vectors, and more similar items are represented by more similar vectors. Hence, STM models that are based on the distinctiveness assumption predict that as similarity increases, order accuracy should decrease, however such models do not predict that the position of the phonemic overlap within the items should be important. Therefore, without the introduction of additional assumptions about the nature of item similarity, these models predict the same detrimental effect on order memory for lists that share the CV_, the _VC or the C_C components.

Most models of STM have been constructed to explain this detrimental effect of similarity on order memory but many seem unable to explain the facilitative effect of rhyme on item memory in the serial recall task reported by Fallon et al. (1999). A possible solution to this problem was offered by Nairne and Kelley (1999). They suggested that the representations of memory items are located in a multidimensional space in which stimuli are represented along a list dimension and a within-list dimension. Sharing phonemic features within a list, but not between lists, has a beneficial effect on item recall because it improves discrimination on the list dimension. Tehan and Fallon (1999) have made a similar suggestion to capture the idea that list cues can be used to facilitate the recall of item information. At the same time, the more similar list items are, the harder it should be to discriminate items on the within-list dimension. That is, to recover an individual item's position within a list. Therefore, according to these models similarity will only increase item recall when the detrimental effect of shared phonological features on the within-list dimension is compensated for by a beneficial effect of similarity on the list dimension (see also Watkins & Watkins, 1975). Nevertheless, such an approach predicts the same facilitative effect on item memory for lists that share the CV_, the _VC or the C_C components, as without additional assumptions each should provide for equally effective discrimination on the list dimension.[6]

To summarise, the non-linguistic models of STM predict that, in comparison to a phonologically dissimilar condition, lists in which any two phonemes are shared by all of the words (CV_, _VC or C_C) will show a detrimental effect on order memory and a facilitative effect on item memory.

---

[6] A unique aspect of the feature model is the assumption that the utility of trace features for recall is not only influenced by the distinctiveness of the list items in relation to other list items, but also by the salience of the trace features. An attentional parameter was implemented into the original feature model to reflect the idea that cues can be used to increase the salience of list items to the extent that participants perceive the items as belonging to the same category.

4.2.1.2. Psycholinguistic models of short-term memory

Some models of verbal STM specifically acknowledge that verbal stimuli may place unique demands on memory and as such any mechanism underlying the recall of verbal material may be quite different from the memory system supporting the recall of other types of stimuli. Any model based on this position could be referred to as psycholinguistic. The strongest form of this argument would be those attempts to model verbal recall solely in terms of processes that have been posited to account for speech perception and production capabilities (e.g., Martin et al., 1999).

One psycholinguistic model of STM (Hartley & Houghton, 1996) that is pertinent to the current experiments has been developed to explain the findings from a small body of research on STM for nonwords. The majority of the research on verbal STM has used words and, as a consequence, most models of STM have focussed on the findings and challenges posed by this literature. However, research on the recall of nonwords poses different challenges for theories of verbal STM. Specifically, recall errors in these studies show a particular pattern that reveals effects of sub-syllabic structure on the recall of nonwords (e.g., Ellis, 1980; Treiman & Danis, 1988). Some researchers (e.g., Gupta & MacWhinney, 1997) have argued that these linguistic constraints must be incorporated into models of STM to provide a full account of performance.

Hartley and Houghton (1996) offer the most developed model of the linguistic constraints on STM that are relevant to understanding the effects of phonological similarity, and any effects of varying the position of the overlapping phonemes within words. This model incorporates two linguistic principles that are crucial to deriving predictions for the following experiments from a psycholinguistic perspective. The first of these is the sonority principle. Sonority refers to the amount of energy in the speech signal and the sonority principle refers to the fact that in syllables sonority increases to a peak in the vowel and then decreases. The strength of a speech trace will not be as strong for consonants as vowels because consonants are shorter in duration and are not as acoustically intense (Hartley & Houghton, 1996). To reflect this, Hartley and Houghton (1996) set the activation level for nodes representing vowels higher than for nodes representing consonants. Hence, this model predicts that any form of similarity will impair order memory, but the greatest impairment will be seen when the vowel is shared, as it is the most strongly represented phoneme in a word.

The second linguistic principle that Hartley and Houghton (1996) have incorporated is that syllables have an internal structure, comprising an onset and a rhyme. The onset consists of the initial consonant or consonant cluster, while the rhyme consists of the vowel and any following consonants. The rhyme is also divided into the peak (the vowel) and the coda (the following consonants). Hartley and Houghton (1996) suggest that syllables are represented by separate nodes corresponding to the onset and the rhyme. At the same time, individual phonemes are associated with slots in a syllable template that serves to maintain the structure of the syllable. Not all of the slots in the template are filled for each syllable, but each occupied slot is also linked to the onset and rhyme nodes. The syllable template preserves the structure of the syllable rather than the content, and is used to activate the onset and rhyme nodes appropriately, which in turn activate individual phonemes. These mechanisms explain why errors in the recall of CVC nonwords are more likely to preserve the VC than the CV or C_C pairings (e.g., Treiman & Danis, 1988), because these errors respect the onset-rhyme structure.

Gupta and MacWhinney (1997) have proposed a conceptual hybrid model by describing how the Hartley and Houghton (1996) model can be incorporated into the Burgess and Hitch (1992) connectionist model of STM. The Burgess and Hitch (1992) model provides a mechanism that can serially order items and explain other recall phenomena, while the Hartley and Houghton (1996) model provides a linguistically constrained model of sub-syllabic processes.

In a hybrid model of this type, the sub-lexical linguistic mechanisms of Hartley and Houghton (1996) can be used to mediate the associations between phonemes and words. As in the Burgess and Hitch (1992, 1999) model, nodes corresponding to words can be associated to a context maintenance queue to maintain the temporal order of the words. If the onset and rhyme nodes are associated with the temporal order mechanism, even indirectly via word units, then this account can explain the facilitative effect of rhyme on item memory. Presentation of rhyming words results in the repeated activation of the same rhyme unit during both presentation and recall, increasing the likelihood of it being correctly recalled.

This psycholinguistic account also predicts that the position of the overlapping phonemes in a list will influence the nature of the phonemic similarity effect observed.

If the words in a list share the CV or the C_C then a different rhyme unit will be activated by each word, just as would occur for a list of dissimilar words, so these lists will not benefit from the repeated activation of the same rhyme unit. Though lists of such items may show some benefit of the repetition of the initial consonant, this effect will be much smaller as the initial consonant is less strongly activated than the vowel (reflecting the sonority principle).

To summarise the psycholinguistic model predicts that, in comparison to lists of dissimilar words, order memory should be reduced for any list where all the items share two phonemes, but this effect will be larger when list items share a common vowel (i.e., _VC, CV_) than when they do not (i.e. C_C). This is because the vowel is the most highly activated phoneme and provides the best discrimination amongst the words. At the same time, item information will be better for lists that share the _VC because the same rhyme unit is repeatedly reinforced, while this is not true for lists that share the CV_ or C_C components.

4.2.2. The current experiments

The following experiments use the same pool of words (sampled without replacement) for all conditions in an experiment, but necessarily use different sets across experiments. The lists are constructed such that the conditions differ across experiments in terms of the position of the shared phonemes within words, yet are equated on phonological similarity. Therefore, in comparison to when the list items are phonemically dissimilar, Experiment 1 will attempt to replicate the item recall advantage that has been found with lists that share a rhyme ending (_VC; Fallon et al., 1999, Experiment 1; Gathercole et al., 1982). The stimulus lists for Experiment 2 will consist of words that share a common initial consonant and vowel (CV_) component, thus changing the position of the overlapping phonemes across experiments, yet keeping the amount of phonemic overlap (as measured by the degree of shared consonant and vowel information) constant. Finally, for Experiment 3 the stimulus lists will consist of words that share common initial and final consonants (C_C).

In each experiment three measures of recall performance will be examined. Correct-in-position refers to the number of items recalled in the position in which they were presented, whereas item recall refers to the number of items recalled regardless of the position in which they were recalled. As performance when scored using the correct-in-position and item recall measures is not independent (Fallon et al., 1999; Murdock, 1976; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d), a measure of order accuracy will also be obtained (i.e., correct-in-position divided by the item recall measure). This yields a measure of the proportion recalled in the correct order as a function of the number of items recalled. Therefore, the better a participant's memory for the order in which the list items were presented, the higher this proportion will be.

These measures provide a test of the predictions derived from the psycholinguistic and non-linguistic models of STM outlined above. To reiterate, non-linguistic models of STM predict that in comparison to phonemically dissimilar items, lists of CV_, _VC or C_C words will have a detrimental effect on order memory when measured using the order accuracy criterion and a facilitative effect on the recall of item information (using the item recall measure). In contrast, psycholinguistic models of STM predict that any form of similarity should impair order memory, but this effect will be largest when the items share a common vowel (i.e., _VC and CV_ lists) as compared to when they do not (i.e., C_C lists). Further, this model predicts an item recall advantage for rhyming lists of items that should be absent (or at least minimal) when list items do not rhyme (i.e., CV_ and C_C lists).

4.3. Experiment 1

4.3.1. Method

4.3.1.1. Participants

Twenty-four undergraduate psychology students from the University of Wollongong participant pool (5 males and 19 females), with an age range from 19 to 46 years (*M* = 23.38), participated in compliance with a course requirement. Only native Australian English speakers who indicated having no prior problems with their hearing participated in the study.

4.3.1.2. Stimuli

The stimuli comprised 180 words with a consonant-vowel-consonant (CVC) phonemic structure (refer Appendix A - Table A1). The stimuli were used to create 30 rhyming, 30 phonemically similar, and 30 phonemically dissimilar six-word lists. Thus, each word was sampled three times, such that each appeared in one rhyming, one similar, and one dissimilar list. For the rhyming condition all of the stimuli in a list shared the _VC component (e.g., *Came, Name, Maim, Lame, Dame,* and *Shame*). For the similar condition two constraints were placed on list construction. The first constraint was that no item in a list shared the _VC component. Also, each stimulus in each list had at least two phonemes in common with at least one other stimulus in the same list (e.g., *Came, Case, Cut, Kip, Cub,* and *Cap*). Therefore, all of the items in a similar list shared the same initial consonant. Finally, for the dissimilar condition, each stimulus in each list did not share any phonemes with any other word in that list (e.g., *Came, Sin, Rang, Leap, Hug,* and *What*).

Using an Arista Cardioid dynamic microphone (Model No. DM-904D), the stimuli were recorded using a Sony Minidisc Deck (Model No. MDS-JE640) in a sound attenuated booth by a female speaker with an Australian English accent. Each stimulus was transferred digitally onto a Macintosh computer and normalised to control for possible amplitude effects on performance. The lists were presented in three blocks of thirty trials. The order of the blocks within the experimental session was counterbalanced across participants. The order of the trials in each block and the order in which the items occurred in each list were randomised for all participants.

4.3.1.3. Procedure

For each condition, two practice lists were given to each participant prior to the presentation of the first experimental list. Each participant was auditorily presented with six words at a rate of one word per second. Stimulus presentation rate was controlled using Hypercard (version 2.4.1). One second after the presentation of the last item in a list, participants heard a 200 ms, 500 Hz tone that was used as a recall prompt. A participant's task was to verbally recall the list items in the order in which they were presented. Participants were told to say 'pass' if they could not remember an item. Thus, strict serial recall instructions were employed. Presentation and recall attempts were recorded onto Minidisc to enable accurate scoring. The time taken for each participant to complete all three conditions was approximately 40 minutes.

4.3.2. Results

The data were analysed using a 3 x 6 (Phonological Similarity x Serial Position) repeated measures analysis of variance (ANOVA) for the correct-in-position measure of performance (i.e., scored as correct if a participant recalled the correct item in the correct position). In addition, item recall (i.e., scored as correct if a participant recalled a list item regardless of position) and order accuracy measures (i.e., correct-in-position divided by the score obtained using the item recall measure) were analysed using two separate repeated measures ANOVAs.[7] Figure 4.1 summarises performance when the stimulus lists were dissimilar, similar or rhyming, collapsed across serial position for each of the three measures. Unless otherwise specified, α was set at .05 (2-tailed). Also, the *Greenhouse-Geisser* statistic is reported instead of the standard *F* statistic where the assumption of sphericity was violated.

The correct-in-position analysis revealed a main effect of phonological similarity, $F(2, 46) = 20.032$, $MSE = 31.426$, $p < .001$. Post hoc paired samples *t*-tests were used to analyse the performance differences across the three conditions. α was set at 0.0167 (2-tailed) to control for the increased probability of committing a Type I error as a function of the number of comparisons performed, thus keeping the family-wise error rate at .05. The analyses revealed that dissimilar lists were recalled more accurately than either similar, $t(23) = 6.609$, $p < .0167$, or rhyming lists, $t(23) = 4.126$, $p < .0167$, which did not differ, $t(23) = 1.391$, *ns.* Although post hoc analyses were not performed on the main effect of position, *Greenhouse-Geisser* (2.691, 61.892) = 181.488, $MSE = 39.577$, $p < .001$, across all three conditions the standard serial position effect with better recall of the initial items and the last item in a list was observed. Phonological similarity was

---

[7] Refer Appendix B for the ANOVA tables for the major statistical analyses performed for study one.

also found to interact with serial position, $F(10, 230) = 5.322$, $MSE = 6.986$, p < .001, such that differences between conditions increased across serial positions.

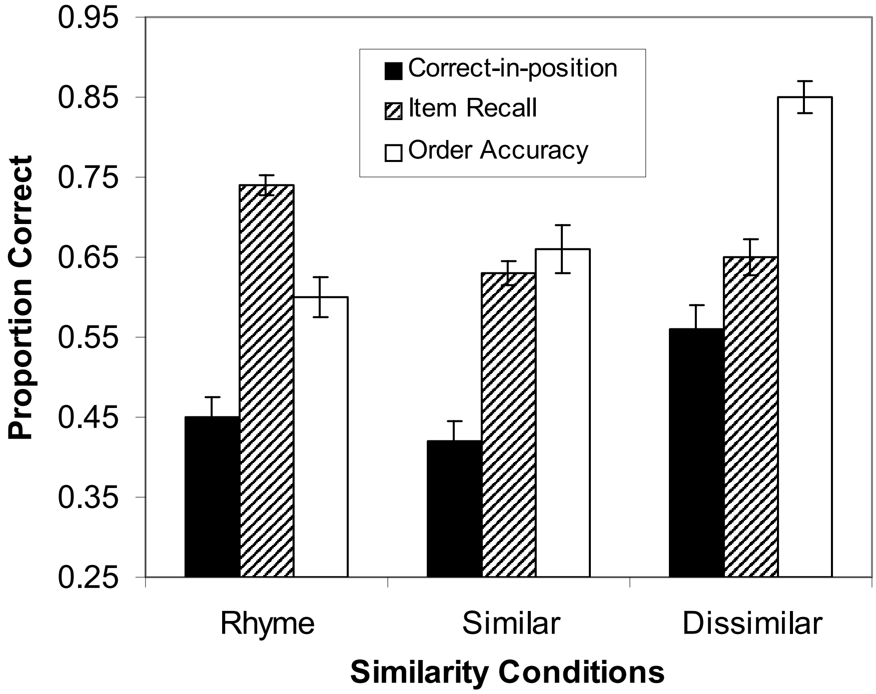

Figure 4.1 Mean proportions correct (± SE) for the phonemically similar, phonemically dissimilar, and rhyming lists of stimuli, for the three scoring procedures (Experiment 1).

The item recall analysis revealed a main effect of phonological similarity, *Greenhouse-Geisser* (1.583, 36.399) = 28.424, $MSE = 123.542$, $p < .001$. Post hoc analyses revealed that item recall was higher for rhyming as compared to either similar, $t(23) = 8.626$, $p < .0167$, or dissimilar lists, $t(23) = 4.885$, $p < .0167$, which did not differ, $t(23) = 1.063$, *ns*. The order accuracy analysis revealed a main effect of phonological similarity, $F(2, 46) = 76.023$, $MSE = .0052$, $p < .001$. Post hoc analyses revealed that dissimilar lists were recalled more accurately than similar lists, $t(23) = 10.316$, $p <.0167$, which were more accurately recalled than rhyming lists, $t(23) = 2.622$, $p < .0167$.

4.3.3. Discussion

Consistent with previous research findings (e.g., Gathercole et al., 1982; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d), regardless of whether performance was scored using the correct-in-position or order accuracy measures, order memory was better for dissimilar as compared to either rhyming or similar lists. Although the correct-in-position criterion yielded no difference in order memory between rhyming and similar lists, when scored using the order accuracy measure, order memory was worse for rhyming as compared to similar lists. In other words, the current research found that order memory was impaired when lists consisted of words that shared a larger number of common phonemes (i.e., rhyming lists) as compared to either similar or dissimilar lists. This finding is consistent with the explanations generated from non-linguistic models of STM (Brown et al., 2000; Burgess & Hitch, 1992; 1999; Nairne, 1988, 1990a; Nairne & Kelley, 1999; Tehan & Fallon, 1999). According to these models as similarity increases order memory should decrease.

Psycholinguistic models of STM can also account for these results. For instance, according to the Hartley and Houghton (1996) model, any form of similarity should impair order memory but the greatest detriment will be seen when the vowel is shared, as is the case for rhyming lists of words.

Consistent with previous research (i.e., Fallon et al., 1999, Experiment 1; Gathercole et al., 1982), the current study found an item recall advantage for rhyming as compared to either similar or dissimilar lists of words. In addition, the same numbers of words were recalled in the similar as compared to dissimilar condition. Therefore, the findings from the current study suggest that not only does the detrimental effect of similarity disappear when performance is measured for item information, but when the stimulus lists rhyme, similarity appears to facilitate the recall of item information whereas a less consistent form of similarity (i.e., similar condition) does not. The idea that the recall of item information is facilitated by retrieval cues is consistent with research that has found that taxonomic category membership can act as a retrieval cue (Huttenlocher & Newcombe, 1976; Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999a). Further, this finding is not only consistent with non-linguistic models which suggest that the rhyme unit acts as a cue to aid in the retrieval of item information

(Nairne, 1988, 1990a, 2002; Nairne & Kelley, 1999; Tehan & Fallon, 1999), but also with the Hartley and Houghton (1996) model in which the rhyme unit serves as a prime to facilitate the recall of item information.

4.4. Experiment 2

Experiment 2 was designed to further examine the explanations generated by STM models for the effect that similarity has on both order and item memory. In one condition of Experiment 2 the stimulus lists consisted of words that shared a common CV_ component. Thus, in comparison to Experiment 1 (i.e., rhyming condition), although the positions of the overlapping phonemes differ, the amount of phonemic overlap has been held constant.

According to non-linguistic models of STM as similarity increases order memory should decrease. The Hartley and Houghton (1996) model also predicts that similarity will impair order memory, however, this effect should be greatest when all of the stimuli in a list share the vowel (i.e., CV_ lists). As such, the predictions generated from STM models to account for the effect that similarity has on order memory for CV_ lists of words are indistinguishable.

The linguistic and non-linguistic models do differ, however, in their predictions regarding the effect that similarity has on the recall of item information. For instance, Nairne and Kelley (1999; see also Nairne, 1988, 1990a; Tehan & Fallon, 1999) suggest that lists that share features that make them easily discriminable along the list dimension should aid in identifying the correct list to recall. Hence, these models would predict an increase in the recall of item information when stimulus lists share the CV_ component.[8] Essentially these models predict that the findings from Experiment 2

[8] While these predictions are correct for the most part, the feature model (Nairne, 1988, 1990a) with the incorporation of an attentional parameter, suggests that it is the salience of the cues that is important for the recall of item information. Hence, it is possible to argue that the rhyme unit may be a more salient cue than other list cues. However, to make a coherent argument, this model would need to specify why this may be the case in comparison to CV_ and C_C lists.

should replicate those observed in Experiment 1, but with CV_ in the place of _VC lists. In contrast, the Hartley and Houghton (1996) model suggests that syllable representations do not get the same degree of support as when the items rhyme because there is not a consistent rhyme unit activated by every item in the list. Hence, this model would predict that the item recall advantage observed when the stimuli rhyme (i.e., Experiment 1) should be absent (or at least minimal) when lists share the CV_ component.

4.4.1. Method

4.4.1.1. Participants

Twenty-four undergraduate psychology students from the University of Wollongong participant pool (4 males and 20 females), with an age range from 18 to 37 years (M = 20.13), participated in compliance with a course requirement. The same inclusion criteria were placed on the selection of participants for the current experiment, as in Experiment 1.

4.4.1.2. Stimuli

The stimuli comprised 180 words with a CVC structure (refer Appendix A - Table A2). The stimuli were used to create 30 six-word lists with a common CV_ component, 30 phonemically similar, and 30 phonemically dissimilar lists. List construction was the same as in Experiment 1 except for a few minor modifications. The first modification was that for the CV_ condition all of the words in a particular list shared the CV_ component (e.g., *Time, Ties, Tight, Type, Tide,* and *Tile*). Also, for the similar condition no item in a list shared the CV_ component (e.g., *Time, Rum, Rhyme, Lime, Limb,* and *Dumb*). Hence, all of the items in a similar list shared the same final consonant.

4.4.1.3. Procedure

The same testing procedure was used in the current experiment as in Experiment 1.

4.4.2. Results

The data were analysed using a 3 x 6 (Phonological Similarity x Serial Position) repeated measures ANOVA for the correct-in-position measure. Also, the scores obtained using the item recall and order accuracy criteria were analysed using two separate repeated measures ANOVAs. Figure 4.2 summarises performance when the stimulus lists were dissimilar, similar or shared the CV_ component, collapsed across serial position for each of the three scoring procedures.

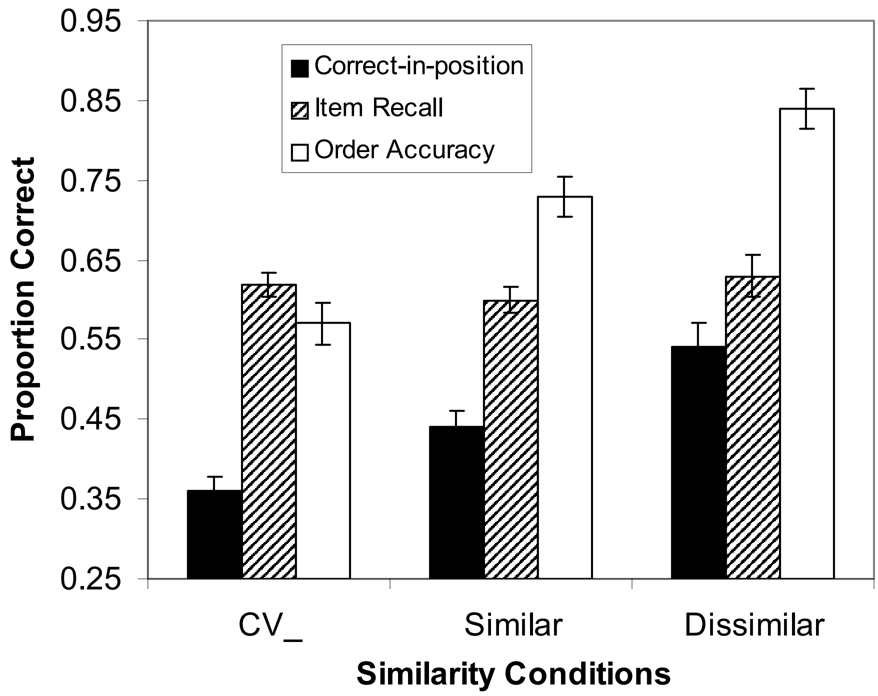

Figure 4.2 Mean proportions correct (± SE) for the phonemically similar, phonemically dissimilar, and CV_ lists for the three scoring procedures.

The correct-in-position analysis revealed a main effect of phonological similarity, $F(2, 46) = 34.996$, $MSE = 30.653$, $p < .001$. Post hoc analyses revealed that dissimilar lists were recalled more accurately than similar lists, $t(23) = 4.876$, $p < .0167$, which were more accurately recalled than CV_ lists, $t(23) = 4.278$, $p < .0167$. A main effect of position was observed, *Greenhouse-Geisser* (3.017, 69.397) = 238.865, *MSE* = 28.210, $p < .001$, and phonological similarity was found to interact with serial position, $F(10, 230) = 4.192$, *MSE* = 6.999, $p < .001$, such that differences between conditions increased across serial positions.

The item recall analysis revealed no phonological similarity effect, *Greenhouse-Geisser* (1.424, 32.745) = 2.517, *MSE* = 138.734, *ns*. The order accuracy analysis revealed a main effect of phonological similarity, $F(2, 46) = 83.061$, *MSE* = .0052, $p < .001$. Post hoc analyses revealed that dissimilar lists were recalled more accurately than similar lists, $t(23) = 6.591$, $p < .0167$, which were more accurately recalled than CV_ lists, $t(23) = 6.659$, $p < .0167$.

4.4.3. Discussion

In terms of the recall of item information, although there was a trend toward a similarity effect, the current study found no differences in item recall levels across the three conditions. On closer inspection this trend was due to a decrease in item recall for the similar as compared to either the dissimilar or CV_ conditions, rather than to an increase in the recall of item information for the CV_ condition. This is important in that Nairne and Kelley (1999; see also Nairne, 1988, 1990a; Tehan & Fallon, 1999) would predict an item recall advantage for stimulus lists that are easily discriminable along the list dimension (i.e., in this case CV_ lists). This prediction is clearly inconsistent with the current research findings. The current findings are, however, consistent with the predictions generated from Hartley and Houghton's (1996) linguistically constrained model of STM. According to this model, an item recall advantage should be observed when list items share a rhyme ending, as was found in Experiment 1. This is because the rhyme unit serves as a prime to reinforce syllable structure. When the phonemic overlap between items does not coincide with sub-syllabic structure (onset/rhyme), as is the case when words share the CV_ component (i.e., Experiment 2), they do not get this additional reinforcement. Hence, in comparison to rhyming lists, the speech traces for CV_ lists of items are less stable. This is consistent with the pattern of results observed in the current study.

In terms of the effect that similarity has on order memory, the predictions generated from both psycholinguistic and non-linguistic models of STM are consistent with the current research findings. For instance, according to non-linguistic models of STM that are based on the distinctiveness assumption (Brown et al., 2000; Burgess & Hitch, 1992; 1999; Nairne, 1988, 1990a; Nairne & Kelley, 1999; Tehan & Fallon, 1999), as similarity increases order memory should decrease. Further, according to the Hartley and Houghton (1996) model any form of similarity should decrease order memory, however, when the vowel is the overlapping phoneme (i.e., in this case, CV_ lists), order memory should be further impaired. As such, these explanations are consistent with the current finding of an order memory impairment for CV_ as compared to similar lists, and for similar as compared to phonemically dissimilar lists.

4.5. Experiment 3

Experiment 3 was again designed to examine the explanations generated by STM models for the effect that phonological similarity has on both order and item memory. The critical stimulus lists for Experiment 3 consisted of words that shared the C_C component. STM models that are based on the distinctiveness argument predict an identical pattern of results to those observed across Experiments 1 and 2. Thus, as similarity increases these models predict a further order memory impairment. Also, STM models that suggest that list cues can be used to facilitate item recall (i.e., Nairne, 1988, 1990a; Nairne & Kelley, 1999; Tehan & Fallon, 1999) predict an item recall advantage for stimulus lists that share the C_C component.

In contrast, although the Hartley and Houghton (1996) model predicts a decrease in order memory for C_C lists of words, because each item in a list has a unique vowel, order memory should not be influenced to the same extent as when the list items either rhyme (i.e., Experiment 1) or share the CV_ component (i.e., Experiment 2). In fact, in comparison to the C_C condition, in this experiment order memory should be poorer for the similar lists because the list items in this condition share a common vowel.

In terms of the recall of item information, the Hartley and Houghton (1996) model predicts an identical pattern of results as in Experiment 2. Thus, the syllable representations will not get the same type of support when list items do not share the rhyme unit. Hence, the item recall advantage observed when stimuli rhyme (i.e., Experiment 1) should be absent (or at least minimal) when the stimuli share the C_C component.

4.5.1. Method

4.5.1.1. Participants

Twenty-four undergraduate psychology students from the University of Wollongong participant pool (2 males and 22 females), with an age range from 18 to 32 years (M = 20.33), participated in compliance with a course requirement. The same inclusion criteria were placed on the selection of participants for the current experiment as for Experiment 1.

4.5.1.2. Stimuli

The stimuli comprised 180 words with a CVC structure (refer Appendix A - Table A3). The stimuli were used to create 30 same-consonant, 30 phonemically similar, and 30 phonemically dissimilar six-word lists. The same constraints were placed on the construction of the stimulus lists as for Experiment 1 with two minor modifications. The first modification was that in the same consonant condition all of the stimuli in a particular list shared the C_C component (e.g., *Bought, Bet, Boot, But, Bat,* and *Bait*). Also, for the similar lists no item in a list shared both consonants (e.g., *Bought, Bored, Lawn, Wrought, Fort,* and *Fawn*). Therefore, all of the items in a similar list shared the same vowel.

4.5.1.3. Procedure

The same testing procedure was used in the current experiment as in Experiment 1.

4.5.2. Results

The data were analysed using a 3 x 6 (Phonological Similarity x Serial Position) repeated measures ANOVA for correct-in-position. In addition, the scored obtained using the item recall and order accuracy measures were analysed using two separate repeated measures ANOVAs. Figure 4.3 summarises performance when the stimulus lists were dissimilar, similar or shared the C_C component, collapsed across serial position for each of the three scoring procedures.

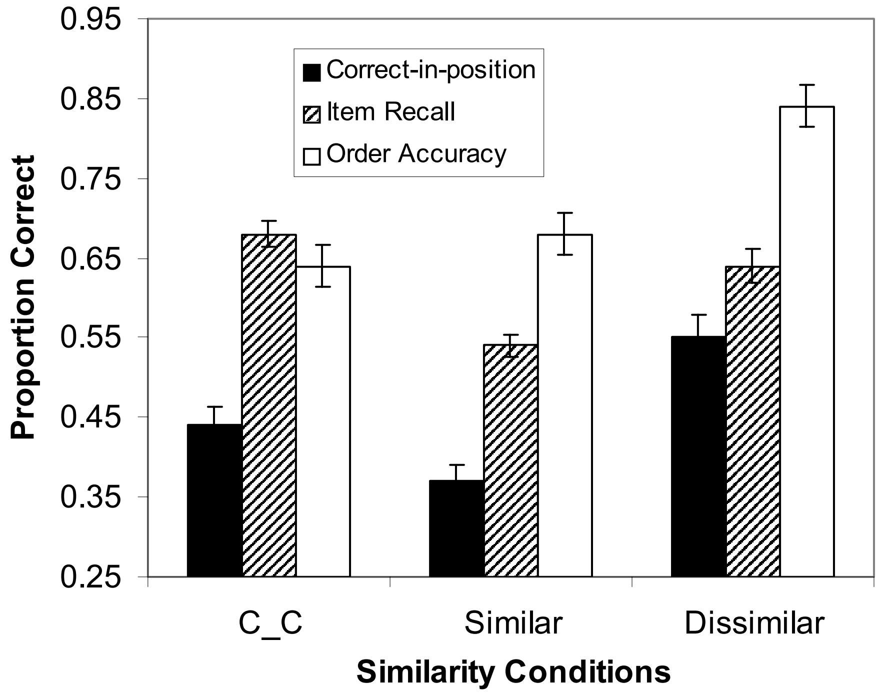

Figure 4.3 Mean proportions correct (± SE) for the phonemically similar, phonemically dissimilar, and C_C Lists for the three scoring procedures.

The correct-in-position analysis revealed a main effect of phonological similarity, $F(2, 46) = 32.172$, $MSE = 31.741$, $p < .001$. Post hoc analyses revealed that dissimilar lists were recalled more accurately than C_C lists, $t(23) = 4.778$, $p < .0167$, which were more accurately recalled than similar lists, $t(23) = 3.384$, $p < .0167$. A main effect of position was observed, *Greenhouse-Geisser* (2.974, 68.410) = 183.313, *MSE* = 29.320, $p < .001$, and phonological similarity was found to interact with serial position, $F(10, 230) = 5.716$, *MSE* = 5.603, $p < .001$, such that differences between conditions increased across serial positions.

The item recall analysis revealed a main effect of phonological similarity, *Greenhouse-Geisser* (1.576, 36.250) = 37.316, *MSE* = 141.030, $p < .001$. Post hoc analyses revealed that in comparison to the phonemically similar condition, item recall was higher for C_C and dissimilar lists, ($t(23) = 11.662$, $p < .0167$; $t(23) = 5.757$, $p < .0167$; respectively). However, the recall of item information did not differ between the C_C and dissimilar lists, $t(23) = 2.043$, *ns*. Finally, the order accuracy analysis revealed a main effect of phonological similarity, $F(2, 46) = 48.578$, *MSE* = .0055, $p < .001$. Post hoc analyses revealed that order memory was better for dissimilar as compared to either C_C, $t(23) = 10.566$, $p < .0167$, or similar lists, $t(23) = 7.314$, $p < .0167$, which did not differ, $t(23) = 1.646$, *ns*.

4.5.3. Discussion

Regardless of whether the correct-in-position or order accuracy criterion was used to measure performance, order memory was better for the dissimilar as compared to either of the phonemically similar conditions (i.e., similar or C_C lists). Thus, the standard detrimental effect of similarity on order memory was observed (e.g., Baddeley, 1966; Baddeley et al., 1984; Gathercole et al., 2001; Lian et al., 2001; Watkins et al., 1974). Although no difference was observed between the C_C and similar conditions when scored using the order accuracy measure, the correct-in-position measure yielded an order memory impairment for similar as compared to C_C lists of words.

This finding of a decrease in order memory (as measured using the correct-in-position criterion) for similar in comparison to C_C lists is inconsistent with explanations generated by non-linguistic models of STM (Brown, et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a; Nairne & Kelley, 1999; Tehan & Fallon, 1999). Furthermore, the finding of no difference in order memory between C_C and similar lists when measured using the order accuracy criterion is also problematic for these models. According to these models as similarity increases order memory should decrease. However, C_C lists share a larger number of overlapping phonemes than do similar lists. Thus, non-linguistic models of STM that suggest that it is the distinctiveness of items in relation to other list items that determines the effect that similarity has on order memory cannot deal with the current research findings.

The results are consistent with psycholinguistic models that suggest that the influence that similarity has on order memory is dependent on the phonemic make up of the experimental list items. For instance, according to the Hartley and Houghton (1996) model, order memory is influenced by both the phonemic similarity of the list items and the effect that sharing a common vowel has on order memory. Therefore, although C_C lists share a greater amount of phonemic overlap than do the phonemically similar lists, the greater weighting of the vowel in comparison to consonants means that order memory is more adversely affected in the similar condition, despite list items in this condition sharing fewer common phonemes.

The findings from the current study regarding the recall of item information are inconsistent with predictions generated from STM models which suggest that shared features that make lists easily discriminable along the list dimension (i.e., C_C lists) aid in the recall of item information (i.e., Nairne, 1988, 1990a, 2002; Nairne & Kelley, 1999; Tehan & Fallon, 1999). If this were the case then the recall of item information for the C_C lists should significantly exceed what was observed for the dissimilar lists and it did not. In contrast, the Hartley and Houghton (1996) model suggests that in comparison to rhyming lists of items, the syllable representations do not get the same level of support when the items share the C_C component. As a consequence, the recall of item information for C_C lists should not benefit to the same extent as when the stimulus lists rhyme.

Although performance for the dissimilar conditions was almost identical across the three experiments, one finding that requires further exploration was the decrease in the recall of item information for the phonemically similar condition in Experiment 3 (i.e., M =.54) as compared to the similar conditions in the other two Experiments (i.e., Experiment 1, M = .63; Experiment 2, M = .60). One possibility is that the differences in item recall levels are due to differences in the prevalence of phoneme recombination errors across these conditions (e.g., the list *sat, bit, sap, map* is recalled as *bat, sit, sap*, *map*). Although recombination errors are rare when the experimental stimuli are familiar and phonemically dissimilar (Hartley & Houghton, 1996), using phonemically similar stimuli may increase the likelihood that participants will produce these types of errors.

To further examine this idea, the proportion of recombination errors that occurred across each of the phonemically similar conditions was collated. Across the phonemically similar conditions, recombination errors made up 7.5 % of the errors in Experiment 1, 11.0 % in Experiment 2, and 15.4 % in Experiment 3. These data were subjected to a one-way ANOVA which revealed a significant difference in the proportion of recombination errors observed across the experiments, $F(2,71) = 12.914$, $MSE = .003$, $p < .001$. Post hoc independent samples *t*-tests revealed that fewer recombination errors occurred when the stimuli consisted of a mixture of C_C and CV_ items (i.e., Experiment 1) as compared to lists that consisted of a mixture of C_C and _VC items (i.e., Experiment 2), $t(46) = 2.696$, $p < .0167$, which in turn produced fewer recombination errors than when the lists consisted of a mixture of CV_ and _VC items (i.e., Experiment 3), $t(46) = 2.672$, $p < .0167$.

This pattern of results is consistent with findings from studies on the types of errors made in speech production and the serial recall of nonwords (Ellis, 1980; MacKay, 1970; Treiman & Danis, 1988) that show that recombination errors are not random. Rather, when a recombination error occurs, research suggests that the error is more likely to occur in the onset (initial phoneme) as compared to coda (final consonant) of a syllable (MacKay, 1970; Treiman & Danis, 1988). In addition, initial consonants tend to substitute for initial consonants and final consonants for final consonants, and vowels are less prone to substitutions than are consonants (Ellis, 1980). To summarise, these findings suggest that recombination errors are more likely to occur from initial phoneme movements and less likely to occur from the movement of vowels. This is consistent with the current research findings in that a lower proportion of recombination errors occurred in the similar condition when the onset was held constant across list items (i.e., Experiment 1) whereas a larger proportion occurred when the vowel was held constant (Experiment 3). Therefore, in comparison to the other phonemically similar conditions, although item recall was lower for the similar condition in Experiment 3, a larger proportion of the errors that occurred in this condition were recombination errors.

It should be noted that this pattern must also reflect the fact that not all recombination errors are detectable. For instance, movement of a phoneme in a list where that phoneme is always the same will result in what appears to be a correct recall. However, the movement of a phoneme that differs across list items may result in what appears to be an order error (e.g., when the initial phoneme in a rhyming list moves). Therefore, recombination errors are most likely to produce a word that was not presented in the list when the list consists of a mixture of CV_ and _VC items (i.e., the phonemically similar condition in Experiment 3).

4.6. General Discussion

Despite the differences in how similarity was operationally defined across experiments, a number of results were consistently observed. For instance, regardless of the way in which order performance was measured (i.e., correct-in-position or order accuracy), order memory was better for dissimilar as compared to any of the similar conditions. This is consistent with earlier research findings of an order memory impairment for phonemically similar lists of stimuli, regardless of the type of stimuli employed, the presentation modality or the recall method used (e.g., Baddeley, 1966; Baddeley et al., 1984; Coltheart, 1993; Cowan et al., 1991; Li et al., 2000; Wickelgren, 1965d). Further, a comparison across Experiments 1, 2 and 3 for the dissimilar conditions revealed similar levels of performance, regardless of whether performance was scored using the item recall (.65, .63, and .64, respectively) or order accuracy (.85, .84, and .84, respectively) measures. This consistency is important in that different stimulus sets were used for each experiment. Hence, if the findings from the current study hinged on stimuli differences, then disparities in the performance measures obtained across the dissimilar conditions would be evident. As such, this finding lends strong support to the suggestion that any differences in item and order memory observed across the current experiments between the dissimilar and similar conditions must be due to the way in which similarity has been operationally defined.

Although commonalities were found across the experiments, an interesting pattern of results emerged in terms of the recall of item information. The first is the findings that in comparison to the dissimilar condition, there was an item recall advantage for stimulus lists that rhymed (Experiment 1). Further, this item recall advantage was absent when the stimulus lists shared the CV_ (Experiment 2) or C_C (Experiment 3) components. Hence, these findings are inconsistent with STM models that suggest that an item recall advantage should be observed whenever lists share features that make them easily discriminable along the list dimension (e.g., Nairne, 1988, 1990a, 2002; Nairne & Kelley, 1999; Tehan & Fallon, 1999). These findings are, however, consistent with the Hartley and Houghton (1996) model. For instance, according to this model the speech trace is more stable when list items share a rhyme ending as compared to when the words in a list do not share this structure (i.e., CV_ and C_C lists).

In terms of the predictions STM models generate for the effect that similarity has on order memory, non-linguistic models suggest that as similarity increases order memory should decrease (Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a; Tehan & Fallon, 1999). It is clear that these models can explain the finding that order memory was better for phonemically similar as compared to either rhyming (Experiment 1) or CV_ lists (Experiment 2), as the overall phonemic overlap between list items is greatest in the latter two conditions. However, what is unclear is how these models could account for the finding that when measured using the correct-in-position criterion, order memory was lower for similar as compared to C_C lists (Experiment 3). In addition, when measured using the order accuracy criterion, no difference in order memory was found between the C_C and similar lists. Both of these findings are problematic for non-linguistic models of STM that assume that it is the distinctiveness of list items in relation to other list items that impairs order memory (Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a; Tehan & Fallon, 1999). This is because the C_C lists shared a greater number of overlapping phonemes than did the similar lists (Experiment 3).

In contrast, the Hartley and Houghton (1996) model suggests that any form of similarity should impair order memory (i.e., C_C lists; Experiment 3). However, when the overlapping phoneme is the vowel, as is the case with rhyming (Experiment 1) or CV_ (Experiment 2) lists, this model predicts a further order memory impairment. Thus, the findings from the current study suggest that the influence that similarity has on order memory is dependent on the phonemic make up of the list items.

In summary, the current research findings suggest that phonological similarity influences both the recall of item information and memory for an item’s position in a list. Also, the current findings rule out the possibility that the item recall advantage observed for rhyming lists of words is due to phonemic overlap. Rather, the results suggest that STM performance is influenced by linguistic mechanisms that operate at the sub-syllabic level. Further, these findings suggest that the influence that similarity has on order memory is dependent on the phonemic make up of the list items in that when similarity is held constant order memory is impaired to a greater extent when the vowel is the overlapping phoneme. Thus, to adequately explain the current research findings it is imperative that STM models incorporate mechanisms that can deal with the psycholinguistic rules that constrain speech production. Hence, the current study has identified an urgent need for STM researchers to integrate linguistic research, and models based on this research, into STM models.

PRELUDE TO CHAPTER 5

Although studies into the effect that phonological similarity has on STM performance have proliferated in recent years, currently only a handful of studies have examined this effect using nonwords (Besner & Davelaar, 1982; Drewnowski, 1980; Fallon et al., in press; Gathercole et al., 2001; Lian et al., 2001). In contrast to the robust finding of an order memory impairment observed in the literature when the stimuli are words (e.g., Conrad & Hull, 1964), when nonwords are used, the findings are contradictory. However, as Wickelgren (1965d) suggests, the standard measure of order memory (i.e., correct-in-position) employed by most STM researchers is influenced by the number of items recalled. Further, the one study that has controlled for the effect that individual differences in item recall ability has on order memory performance (i.e., Fallon et al., in press) found the standard detrimental effect in that order memory was worse for phonemically similar as compared to dissimilar lists of nonwords.

Currently, the majority of STM models do not provide an account for the effect that similarity has on STM performance when the stimuli are nonwords. This stems from the belief that, "…given that no adequate long-term representations are available for nonwords, the reconstruction process, for all practical purposes, is thought not to operate for these items" (Saint-Aubin & Poirier, 2000; p.333). Those models that do provide an explanation (e.g., Nairne, 1988, 1990a, 2002; Gupta & MacWhinney, 1997; Hartley & Houghton, 1996), differ with respect to the mechanisms they assume are responsible for this effect. As a consequence, the predictions that STM models generate for the effect that the phonemic similarity of nonword has on STM performance differ.

There were three aims for conducting the nonword experiments. The first aim was to replicate the Fallon et al. (in press) finding of a detrimental PSE on order memory for nonwords. Secondly, to examine whether, as is the case when the stimuli are words (Study 1), operationally defining phonemic similarity in different ways, differentially influences item and order memory when the stimuli are nonwords. Finally, the current experiments were designed to examine the predictions that non-linguistic (e.g., Nairne, 1988, 1990a, 2002) and psycholinguistic (e.g., Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) models of STM generate for the effect that similarity has on STM performance when the experimental stimuli are nonwords.

5. INVESTIGATING THE PHONOLOGICAL SIMILARITY EFFECT: ARE NONWORDS AN ALIEN FORM?*

By

Lisa M. Nimmo and Steven Roodenrys

University of Wollongong
Wollongong, Australia

Address for correspondence:
Lisa Nimmo, Department of Psychology, University of Wollongong
Northfields Ave, Wollongong, 2522, Australia
Phone: (+61) 242 214 454, Fax: (+61) 242 214 163
Email: lmn02@uow.edu.au

Running head: Phonological similarity, nonwords, and serial recall

*This paper is currently under review (Journal of Experimental Psychology: Learning, Memory, and Cognition)

5.1. Abstract

The current research examined the predictions that short-term memory (STM) models generate for the phonological similarity effect, when similarity was defined in different ways. Three serial recall experiments with consonant-vowel-consonant (CVC) nonwords are reported, where the position of the phonemes that list items shared was manipulated (i.e., list items shared the vowel and final consonant [_VC; Experiment 1], the initial consonant and vowel [CV_; Experiment 2], or the two consonants [C_C; Experiment 3]). The results show that the position of common phonemes in nonwords has differential effects on order and item information. The findings are discussed in relation to previous research into the effect that phonemic similarity has on nonword recall, and modifications to current STM models are proposed.

5.2. Introduction

One of the most prominent and robust findings in the research literature on verbal short-term memory (STM) is the phonological similarity effect: the finding of poorer serial recall of lists of words that sound similar to each other as compared to lists of distinct sounding words. Numerous studies of this effect have led to the view that phonological similarity predominantly disrupts memory for the order of the words in the list rather than memory for the identity of the words (e.g., Baddeley et al., 1984; Coltheart, 1993; Conrad et al., 1966; Conrad & Hull, 1964; Cowan et al., 1991; Laughery & Pinkus, 1966; Schweickert et al., 1990).

Although the phonological similarity effect on the serial recall of word lists is a benchmark finding in the research literature on STM, only a handful of studies have explored the effect that phonological similarity has on STM performance with lists of nonwords (Besner & Davelaar, 1982; Drewnowski, 1980; Fallon et al., in press; Gathercole et al., 2001; Lian et al., 2001). In stark contrast to research findings when the stimuli are words, the results of the small number of studies using lists of nonwords are contradictory. The aim of the experiments reported below was to clarify the nature of the phonological similarity effect on nonword recall by examining the effect that different operational definitions of phonological similarity has on STM performance, and to test the predictions derived from a number of STM models.

Before discussing the research on phonological similarity and nonword recall it is worth noting that in several models of serial recall the phonological similarity effect is related to another effect, that of lexicality. The lexicality effect refers to the finding that recall performance is superior for words as compared to when the stimuli are nonwords (e.g. Hulme et al., 1991). A number of STM models (e.g. Brown & Hulme, 1995; Schweickert, 1993) incorporate a process termed ‘redintegration’ in which, prior to output, partially degraded traces held in STM are reconstructed by comparing them to representations stored in long-term memory (akin to a “clean-up” process in connectionist models). According to this view, memory span for nonwords is lower because there are no stored representations available to assist in the reconstruction of a partial trace (Hulme et al., 1997). Significantly for the present research, in a number of models the phonological similarity effect for words also arises in the redintegration process (e.g. Brown et al., 2000; Burgess & Hitch, 1992, 1999; Page & Norris, 1998). A degraded trace retrieved from STM is more likely to be incorrectly identified as a similar word from the list than a dissimilar word, creating an order error. If nonwords lack a representation in long-term memory it is unclear how these models would explain a phonological similarity effect in nonword recall.

Nevertheless, a phonological similarity effect has been found in the recall of lists of nonwords, although the pattern of results is inconsistent. For instance, Lian et al. (2001) examined the recall of sets of nonwords that differed in associative value – nonwords that are high associative value are more wordlike. They compared performance on dissimilar, phonemically similar, and rhyming lists of nonwords and found that order memory was better for dissimilar lists of nonwords that were rated as being high in associative value. However, when the stimuli were rated as being low in associative value, they found no effect of similarity for rhyming lists of nonwords in one experiment (Experiment 1B), and an advantage for the recall of phonemically similar lists[9] in another experiment (Experiment 1A). Lian et al. (2001) suggest that nonwords that are classified as being high in associative value are more similar to real words than those rated low in associative value. According to this view, the more wordlike an item is, the easier it is to access lexical representations held in LTM. These lexical representations can then be used to aid in the retrieval process. This idea is consistent with previous research that suggests that the more wordlike a nonword is rated, the more accurately the nonword is recalled (e.g., Gathercole, 1995; Gathercole et al., 1991b; Gathercole & Martin, 1996; Metsala, 1999).

However, there is an alternative suggestion that may account for the inconsistencies observed across the Lian et al. (2001) study. When the serial recall task requires participants to verbally produce the presented list items, performance may be influenced by an individual's articulatory ability (see, Gathercole et al., 1999; Snowling

---

[9] Please note, Lian et al. (2001; Experiment1A) did not provide enough detail as to how the phonemically similar stimulus sets were constructed (i.e., all sharing a vowel, sharing initial or final consonants, or a mixture of phonemes in different positions). This is important in that Fallon et al. (1999) suggest that differential results emerge in the literature depending on how similarity has been operationally defined.

& Hulme, 1989; Wells, 1995). It may be that saying a list of either rhyming (e.g., *wut, vut, zut, yut, chut*) or phonemically similar (e.g., *zut, fub, zun, zug, fup*) nonwords is easier than saying a list of phonemically dissimilar nonwords (e.g., *zut, yied, hig, chone, wabe*). Furthermore, as Wickelgren (1965d) suggests, the correct-in-position measure of order memory (i.e., scored as correct if a participant recalls the correct item in the correct position) is influenced by the total number of items recalled irrespective of position. Hence, the findings observed when performance is measured using the correct-in-position criterion may be influenced by other factors, such as articulatory ease (see Murray & Jones, 2002), that may act to constrain the number of items an individual is able to recall.

To date, Fallon et al. (in press) have conducted the only study of the phonemic similarity effect with nonwords that has measured item memory, and used a measure of order memory that takes into account individual differences in item recall ability (correct-in-position divided by the score obtained using the item recall measure – termed “order accuracy”). Consistent with the findings from a number of word studies (e.g., Fallon et al., 1999; Gathercole et al., 1982), Fallon et al. (in press) found an item recall advantage for lists of rhyming nonwords. In terms of the effect of phonemic similarity on order memory, they found that correct-in-position recall was better for rhyming as compared to dissimilar lists. However, when scored using the order accuracy criterion, similarity was found to impair order memory. In other words, after controlling for individual differences in item recall ability, order memory was better for dissimilar as compared to rhyming lists of nonwords. In summary, consistent with the findings observed when the experimental lists are words (e.g., Fallon et al., 1999), the results from nonword studies suggest that phonological similarity has differential effects on item and order memory.

5.2.1. Short-term memory models and the phonological similarity effect

Given the robust nature of the phonological similarity effect, most STM models incorporate mechanisms to account for it (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Gupta & MacWhinney, 1997; Nairne, 1988, 1990a; Tehan & Fallon, 1999).

However, as mentioned above, to date the majority of STM models do not provide an explanation for the effect that the phonemic similarity of nonwords has on STM performance. This stems from the belief that, "…given that no adequate long-term representations are available for nonwords, the reconstruction process, for all practical purposes, is thought not to operate for these items" (Saint-Aubin & Poirier, 2000; p.333). Those that do provide an explanation for the effect that similarity has on the recall of nonwords can be divided into two classes: psycholinguistic and non-linguistic models of STM. Psycholinguistic models of STM (e.g., Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) attribute the effect that phonological similarity has on STM to linguistic constraints that are assumed to operate at the sub-syllabic level. In contrast, non-linguistic models of STM (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a) are based on the ratio rule, or idea that the likelihood of recalling a presented list item is relative to the phonemic similarity of all of the presented list items (Gillund & Shiffrin, 1984; Hintzman, 1986; Luce, 1959; Nosofsky, 1986). Although psycholinguistic models, as well as a number of non-linguistic models of STM, argue that phonological similarity differentially influences item and order memory, the underlying mechanisms that they assume are responsible for this effect differ. Hence, we will present a brief description of the two classes of models, with an emphasis on the predictions that each class of model generates for the effect that similarity has on both item and order memory.

5.2.1.1. Psycholinguistic models of short-term memory

Psycholinguistic models of STM were designed to account for linguistic research findings that suggest that sub-syllabic structures influence the recall of nonwords (see Ellis, 1980; Treiman & Danis, 1988). Hartley and Houghton (1996) have developed one of the most detailed STM models of this type. Their linguistically constrained model of STM is based on two linguistic principles: syllable structure and sonority. Linguistic research suggests that syllables have an internal structure comprising an onset and a rhyme (Fudge, 1969; Goswami & Bryant, 1990; Treiman, 1983, 1986; Treiman & Zukowski, 1990). The onset consists of the initial consonant or consonant cluster, whereas the rhyme includes the vowel and subsequent consonants. To model the effect that syllable structure has on STM performance, Hartley and Houghton (1996) have incorporated a syllable layer into their model. At this level in the model, two separate nodes, corresponding to the onset and rhyme are used to represent each syllable.

Sonority refers to the energy of a speech trace, and the sonority principle, to the fact that the sonority of a syllable increases to a peak at the vowel and then decreases. Hartley and Houghton (1996) set the activation level of vowels higher than consonants to reflect the idea that vowels are both longer in duration and more acoustically intense. This leads to the prediction that any form of similarity can potentially disrupt the recall of order information, however, this effect will be larger when items share the vowel.

The detailed specification of the Hartley and Houghton (1996) model allows specific predictions to be derived for the effect of phonemic similarity on both item and order information when the nature of the phonemic overlap between items varies. This model can explain the beneficial effect on item information when lists rhyme (e.g. Fallon et al., in press) because consonant-vowel-consonant (CVC) items share the same rhyme unit (_VC). Repeated activation of the same rhyme unit serves as a prime to facilitate the recall of item information. However, when list items share two phonemes but do not rhyme (i.e., consonant-vowel [CV_] or consonant-consonant [C_C]), syllable representations do not receive the same level of support. Therefore, item recall should not benefit to the same extent for lists of these items.

In terms of the effect that phonemic similarity has on order memory, the Hartley and Houghton model (1996) predicts that, although any form of similarity should decrease a participant's ability to recall items in the correct order, the greatest impairment will be seen when the overlapping phoneme is the vowel (i.e., lists of items sharing the _VC and CV_) because it is more strongly weighted.

5.2.1.2. Non-linguistic models of short-term memory

In contrast to models designed to incorporate language based performance constraints on STM performance, non-linguistic models of STM are based on general principles, without regard for stimulus type (e.g., pictures or spatial location). Nairne's (1988, 1990a) feature model of immediate memory can be used as an exemplar of what is meant by a non-linguistic STM model. Like the psycholinguistic model described above, the feature model (Nairne, 1988, 1990a, 2002; Neath, 1999) suggests that phonological similarity differentially influences item and order memory.

According to the feature model (e.g., Nairne, 2002; Neath, 1999) both short-term and long-term representations are formed during the experimental session, regardless of whether the stimuli are words or nonwords.[10] In other words, the long-term (secondary memory) representations are context specific. At recall, order of recall is determined by the retrieval of degraded traces from a short-term store, however, those traces are identified by comparison against traces in long-term memory, leading to both item and order errors. In terms of the recall of item information, Nairne (2002) suggests that phonemic similarity should facilitate item recall when list items share unique features that can be used as retrieval cues to limit the size of the 'secondary memory search set' (e.g., a common rhyme ending). Hence, this model predicts an item recall advantage when list items share phonemes that can be used as retrieval cues, so lists sharing the _VC, CV_ or C_C components should all be equally well recalled. In terms of the effect that similarity has on order memory, as with other non-linguistic models of STM (Brown et al., 2000), the feature model is based on the ratio rule or the distinctiveness assumption. Hence, according to these types of models, as similarity increases order memory should decrease.

5.2.2. The current experiments

The current work aimed to test the predictions derived above from two types of STM models regarding the effect that phonemic similarity has on both order and item memory in nonword recall. Across experiments, lists of nonwords were constructed that shared the same amount of phonemic overlap, but differed with respect to the position of the shared phonemes. For instance, in comparison to phonemically dissimilar lists of nonwords, the stimulus lists used in the current study shared the _VC (Experiment 1), the CV_ (Experiment 2) or the C_C (Experiment 3) components. Further, as the correct-in-position measure of performance is not independent of a participant's item recall

---

[10] Nairne (2002) does not differentiate between word and nonword performance. As such, it is unclear how current versions of the feature model could explain the lexicality effect.

ability (Fallon et al., 1999; Murdock, 1976; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d), a measure of order accuracy was obtained. This yields a measure of the proportion correct as a function of the number of items recalled.

As outlined previously, both non-linguistic and psycholinguistic models of STM generate predictions for the effect that similarity has on the recall of item information, and memory for an item's position in a list. However, these predictions differ as a consequence of the mechanisms that each class of models has implemented to account for this effect. For instance, in terms of the recall of item information (as measured using the item recall criterion), the feature model (Nairne, 2002) predicts an item recall advantage for lists that share unique features (i.e., in this case, _VC, CV_, and C_C lists). The Hartley and Houghton (1996) model also predicts an item recall advantage for rhyming lists of nonwords, in that the rhyme unit serves as a prime to facilitate the recall of item information. However, when list items do not rhyme (i.e., CV_ lists in Experiment 2 and C_C lists in Experiment 3), item recall should not benefit to the same extent because the speech traces held in STM do not receive the same level of support.

In terms of the effect that phonemic similarity has on order memory (as measured using the order accuracy criterion), the feature model (Nairne, 1988, 1990a, 2002) predicts that, as similarity increases order memory should decrease. Hence, the greater the phonemic overlap between list items (i.e., _VC, CV_ and C_C lists), as compared to either phonemically similar or dissimilar lists of nonwords, the worse a participant's memory for an item's position in a list. In contrast, the Hartley and Houghton (1996) model has been modified to reflect the idea that the speech trace is stronger for vowels as compared to consonants. Thus, according to this model, changing the position of the shared phonemes within list items while holding the number of shared phonemes constant should differentially influence order memory. In other words, order memory should be lower for list items that share a common vowel (i.e., _VC and CV_ lists) in comparison to the similar or dissimilar conditions. This can be contrasted with order memory when each item in a list has a distinct vowel (i.e., C_C lists).

5.3. Experiment 1

5.3.1. Method

5.3.1.1. Participants

Twenty-four undergraduate psychology students from the University of Wollongong participant pool (5 males and 19 females), with an age range from 18 to 56 years (*M* = 23.21), participated in compliance with a course requirement. Only native Australian English speakers who indicated having no prior problems with hearing participated in the study.

5.3.1.2. Stimuli

The stimuli comprised 150 nonwords/nonsense words with a consonant-vowel-consonant (CVC) phonemic structure (refer Appendix C – Table C1). The stimuli were used to create 30 rhyming, 30 phonemically similar, and 30 phonemically dissimilar five-item lists. Thus, each nonword was sampled three times, such that each appeared in one rhyming, one similar, and one dissimilar list. For the rhyming condition, all of the stimuli in a particular list shared the _VC component (e.g., *Vame, Pame, Yame, Wame,* and *Zame*). For the similar condition, two constraints were placed on list construction. The first constraint was that no item in a list shared the _VC component. Also, each stimulus in each list shared at least two phonemes with at least one other stimulus in the same list (e.g., *Pame, Pone, Pog, Pome,* and *Pag*). Therefore, all of the items in a similar list shared the same initial consonant. Finally, for the dissimilar condition, each stimulus in each list did not share any phonemes with any other stimulus in that list (e.g., *Pame, Lun, Teeb, Hoke,* and *Vag*).

Using an Arista Cardioid dynamic microphone (Model No. DM-904D), the stimuli were recorded using a Sony Minidisc Deck (Model No. MDS-JE640) in a sound attenuated booth by a female speaker with an Australian English accent. Each stimulus was transferred digitally onto a Macintosh computer and normalised to control for possible amplitude effects on performance. Before testing began, five participants that did not take part in the experiment were asked to listen to, and repeat each nonword to check their audibility. If more than one participant repeated the same nonword incorrectly, the nonword was re-recorded and another five participants were asked to listen to, and repeat each nonword. The criterion (i.e., no more than one participant repeated the same nonword incorrectly) for satisfactory audibility of the nonwords was met. Overall, participants correctly repeated 99.2% of the nonwords presented. The lists were presented in three blocks of thirty trials. The order of the blocks within the experimental session was counterbalanced across participants. The order of the trials in each block, and the order in which the items occurred in each list, were randomised for all participants.

5.3.1.3. Procedure

Across all conditions, two practice lists were given to each participant prior to the presentation of the first experimental list. Each participant was auditorily presented with five nonwords at a rate of one nonword per second. Stimulus presentation rate was controlled using Hypercard (version 2.4.1). One second after the presentation of the last item in a list, participants heard a 200 millisecond, 500 Hz tone that was used as a recall prompt. The participant's task was to verbally recall the list items in the order in which they were presented. Participants were told to say 'pass' if they could not remember an item. Thus, strict serial recall instructions were employed. Presentation and recall attempts were recorded onto Minidisc to enable accurate scoring. The recordings from three randomly selected participants were transcribed and scored independently by a researcher who was familiar with the scoring rules used in the current study. Inter-rater reliability scores for the rhyming (correct-in-position recall, 98%; item recall, 100%), similar (correct-in-position recall, 99%; item recall, 97%) and dissimilar (correct-in-position recall, 99%; item recall, 99%) conditions were obtained. The time taken for each participant to complete all three conditions was approximately 40 minutes.

5.3.2. Results

Traditionally, performance across serial positions is examined using the correct-in-position scoring criterion. However, the measure of performance obtained when using this criterion is not independent of a participant's item recall ability (Fallon et al., 1999; Murdock, 1976; Poirier & Saint-Aubin, 1996; Wickelgren, 1965d). Therefore, the current study obtained a measure of correct-in-position recall purely to make the current findings directly comparable with other research into the effect that phonemic similarity has on STM performance. Further, obtaining a measure of correct-in-position performance is a necessary step for calculating the order accuracy measure (i.e., correct-in-position divided by the score obtained using the item recall measure). The correct-in-position (i.e., scored as correct if a participant recalled an item in the correct position), item recall (i.e., scored as correct if a participant recalled an item presented in a given list, regardless of position) and order accuracy measures were analysed using three separate repeated measures analyses of variance (ANOVAs).[11] Figure 5.1 summarises performance for the dissimilar, similar, and rhyming lists of nonwords, collapsed across serial position for each of the three performance measures. Unless otherwise specified, α was set at .05 (2-tailed). Also, the *Greenhouse-Geisser* statistic was quoted instead of the standard *F* statistic where the assumption of sphericity was violated.

The correct-in-position analysis revealed a main effect of phonological similarity, $F(2,46) = 16.672$, $MSE = 97.516$, $p < .001$. Post hoc paired samples *t*-tests were used to analyse the differences in recall performance across the three conditions. α was set at 0.0167 (2-tailed) to control for the increased probability of committing a Type I error as a function of the number of comparisons performed, thus keeping the family-wise error rate at .05. Post hoc analyses revealed that dissimilar lists were recalled less accurately than either of the similar conditions, (rhyming, $t(23) = 5.041$, $p < .0167$; similar, $t(23) = 4.591$, $p < .0167$), which did not differ, $t(23) = 1.250$, *ns*.

The item recall analysis revealed a main effect of phonological similarity, $F(2, 46) = 118.636$, $MSE = 89.092$, $p < .001$. Post hoc analyses revealed that item recall was higher for rhyming as compared to similar lists, $t(23) = 7.123$, $p < .0167$, which was higher than the dissimilar lists, $t(23) = 9.055$, $p < .0167$. The order accuracy analysis revealed a main effect of phonological similarity, $F(2,46) = 85.938$, $MSE = .0042$, $p < .001$. Post hoc analyses revealed that dissimilar lists were recalled more accurately than similar lists, $t(23) = 8.439$, $p <.0167$, which were more accurately recalled than when the stimuli rhymed, $t(23) = 5.334$, $p < .0167$.

---

[11] Refer Appendix D for the major for the ANOVA tables for the major statistical analyses performed for study two.

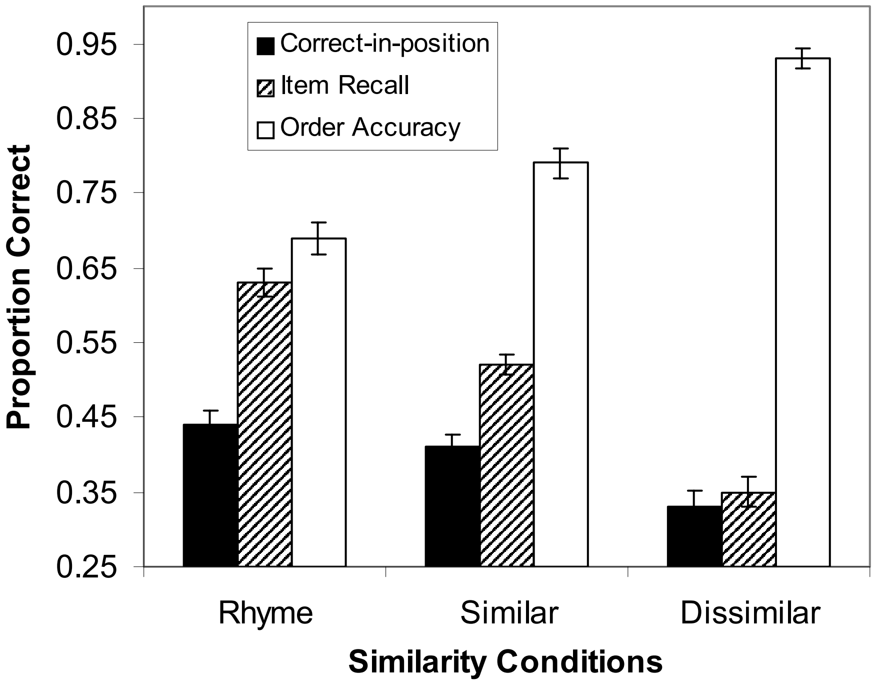

Figure 5.1 Mean proportions correct (± *SE*) for the similar, dissimilar, and rhyming lists of nonwords for the three scoring procedures (Experiment 1).

5.3.3. Discussion

When scored using the correct-in-position criterion, order memory was better for nonword lists that shared some form of similarity (regardless of whether they rhymed or were phonemically similar). This finding of an order memory advantage for phonemically similar as compared to dissimilar lists is inconsistent with previous research which suggest that the locus of the phonological similarity effect lies in the detrimental effect that similarity has on order memory, regardless of whether the stimuli are words (Baddeley, 1966; Farrell & Lewandowky, 2003; Li et al., 2000; Poirier & Saint-Aubin, 1996; Watkins et al., 1974) or nonwords (Besner & Davelaar, 1982; Gathercole et al., 2001). As suggested previously, however, the measure of order memory obtained when using the correct-in-position criterion is not independent of differences in item recall levels across conditions (Wickelgren, 1965d).

When the influence that individual differences in item recall has on the number of possible order errors was controlled (using the order accuracy measure), the current study found that order memory was better for dissimilar as compared to the phonemically similar lists, which was better than when the nonword lists rhymed (see Study 1, for an identical pattern of results with words). Further, this finding of a detrimental effect of similarity on order memory for nonwords is consistent with a recent study conducted by Fallon et al (in press). This is important in that Fallon et al. (in press) have conducted the only nonword study to date that has used the order accuracy criterion to examine the effect that similarity has on order memory.

The finding that similarity impairs order memory for lists of nonwords is also consistent with the predictions generated from non-linguistic, as well as psycholinguistic models of STM. For instance, non-linguistic STM models that are based on the ratio rule (e.g., Nairne, 2002), predict that as similarity increases order memory should decrease. Further, the Hartley and Houghton (1996) model suggests that any form of similarity should impair order memory (e.g., phonemically similar as compared to the dissimilar lists), however, when the overlapping phoneme is the vowel (i.e., as is the case when list items rhyme), order memory should be further impaired.

In terms of the effect that similarity has on the recall of item information, the current study found an item recall advantage for rhyming as compared to phonemically similar lists, which produced more items than when the nonwords were phonemically dissimilar. This finding is consistent with recent research that has found an item recall advantage for rhyming lists, regardless of whether the stimuli are words (e.g., Fallon et al., 1999, Experiment 1; Gathercole et al., 1982; see also Study 1) or nonwords (e.g., Fallon et al., in press).

Furthermore, both the feature model (Nairne, 2002) and Hartley and Houghton's (1996) linguistically constrained model of STM can provide plausible explanations for the item recall advantage observed in the current study. According to the feature model, an item recall advantage should be observed whenever list items share unique features, such as a common rhyme ending, that can be used to limit the size of the memory search set (Nairne, 2002). In contrast, the Hartley and Houghton (1996) model represents syllables in terms of items with a distinct onset-rhyme structure. When each item in a list shares a common rhyme unit, this structure can be used to reinforce the syllable representations. The result of this process is an item recall advantage for rhyming lists of nonwords.

To summarise, when stimuli rhyme, both the feature model (Nairne, 2002) and Hartley and Houghton's (1996) linguistically constrained model of STM can provide plausible explanations for the effect that phonemic similarity has on both the recall of item information, and memory for an item's position in a list. Hence, two further experiments are reported that were designed to distinguish between these two competing explanations. This is important in that the findings from Study 1 suggest that the effect that similarity has on item and order memory is not a consequence of the degree to which list items share common features. Rather, the findings suggested that similarity either facilitates or has a detrimental effect on STM performance as a result of linguistic mechanisms that operate at the sub-syllabic level. Therefore, in Experiment 2 the important condition consisted of lists of nonwords that shared the CV_ component, whereas for Experiment 3, the important condition consisted of nonword lists that shared the C_C component.

5.4. Experiment 2

According to Nairne (2002), a beneficial effect of similarity for the recall of item information should be observed whenever the size of the secondary memory search set can be limited to a smaller set of possible items. Hence, if the item recall advantage observed for rhyming lists of nonwords is due to limiting the size of this search set, then an item recall advantage should also be observed for other lists where the search set is limited to a smaller number of items (i.e., in this case, items that share the CV_ component).[12] In contrast, the Hartley and Houghton (1996) model suggests that when

---

[12] The feature model (Nairne, 1988, 1990a) includes an attentional parameter to reflect the idea that the recall of item information is influenced not only by the distinctiveness of list items in relation to the other presented list items, but also by the salience of the cues. Therefore, it is possible to argue that the rhyme unit is a more salient cue than other types of list cues. However, to make a logical argument, this model would need to specify why this may be the case as compared the CV_ lists.

list items do not share the rhyme structure, syllable representations do not receive the same level of support. Therefore, according to this model, the recall of item information should not benefit to the same extent when list items share a CV_ component. This is not to suggest that an item recall advantage for CV_ lists of nonwords is not possible, but merely that sharing a common rhyme ending should facilitate the recall of item information to a greater extent than when lists share other phoneme pairs.

Furthermore, although the predictions that each model generates for the effect that similarity has on order memory for Experiment 2 are identical, the mechanisms that they assume are responsible for this order memory impairment are not. For instance, according to STM models that are based on the distinctiveness assumption, as similarity increases order memory should decrease. Thus, non-linguistic models of STM predict an identical pattern of results to those observed in Experiment 1 (i.e., order memory should be better for dissimilar as compared to similar lists, which in turn should be better than lists of CV_ nonwords). Further, the Hartley and Houghton (1996) model suggests that although similarity should have a detrimental effect on order memory (i.e., phonemically similar as compared to dissimilar lists), the greatest impairment should be observed when the overlapping phoneme is the vowel (i.e., CV_ lists in the current experiment). Therefore, although both types of models make the same predictions regarding the effect that phonemic similarity has on order recall, crucially these models differ in their predictions for the recall of item information.

5.4.1. Method

5.4.1.1. Participants

Twenty-four undergraduate psychology students from the University of Wollongong participant pool (5 males and 19 females), with an age range from 16 to 49 years (M = 21.54), participated in compliance with a course requirement. As in Experiment 1, the same inclusion criteria were placed on the selection of participants.

5.4.1.2. Stimuli

The stimuli comprised 150 nonwords with a CVC structure (refer Appendix C – Table C2). The stimuli were used to create 30 same initial consonant and vowel (CV_), 30 phonemically similar, and 30 phonemically dissimilar five-item lists. List construction was the same as in Experiment 1, except for two minor modifications. The first modification was that for the CV_ condition, all of stimuli in a particular list shared the CV_ component (e.g., *Maib, Maip, Maig, Maif,* and *Maich*). Also, for the similar condition, no item in a list shared the CV_ component (e.g., *Maib, Mab, Wab, Raib,* and *Wieb*). Hence, all of the items in a similar list shared the same final consonant.

5.4.1.3. Procedure

The same testing procedure was used in the current experiment as in Experiment 1. The criterion for satisfactory audibility of the nonwords was met with participants correctly repeating 99.5% of the nonwords presented. Inter-rater reliability scores for the CV_ (correct-in-position recall, 99%; item recall, 100%), similar (correct-in-position recall, 99%; item recall, 99%) and dissimilar (correct-in-position recall, 97%; item recall, 96%) conditions were obtained.

5.4.2. Results

The correct-in-position, item recall, and order accuracy measures were analysed using three separate repeated measures ANOVAs. Figure 5.2 summarises performance for the dissimilar, similar, and CV_ conditions, collapsed across serial position for each of the three performance measures.

The correct-in-position analysis revealed a main effect of phonological similarity, $F(2,46) = 40.603$, $MSE = 62.246$, $p < .001$. Post hoc analyses revealed that similar lists were recalled more accurately than either CV_, $t(23) = 8.368$, $p < .0167$, or dissimilar lists, $t(23) = 6.456$, $p < .0167$, which did not differ, $t(23) = .752$, *ns.* The item recall analysis revealed a main effect of phonological similarity, $F(2,46) = 83.751$, $MSE = 64.212$, $p < .001$. Post hoc *t*-tests revealed that similar lists were recalled more accurately than CV_ lists, $t(23) = 2.744$, $p < .0167$, which were more accurately recalled than dissimilar lists, $t(23) = 9.788$, $p < .0167$. The order accuracy analysis revealed a main effect of phonological similarity, $F(2,46) = 95.719$, $MSE = .0065$, $p < .001$. Post hoc analyses revealed that dissimilar lists were recalled more accurately than similar lists, $t(23) = 5.358$, $p < .0167$, which were more accurately recalled than CV_ lists, $t(23) = 8.833$, $p < .0167$.

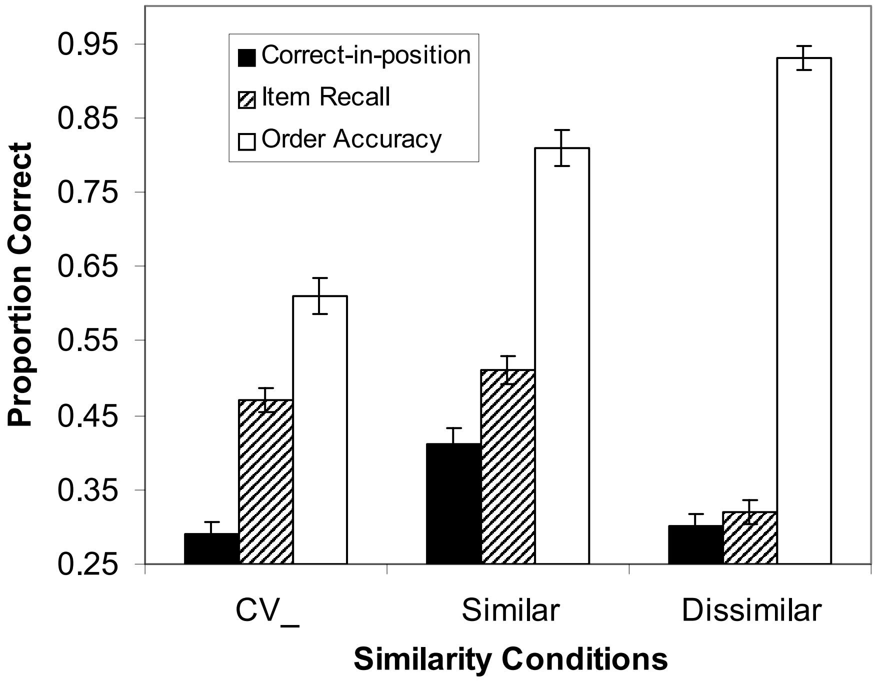

Figure 5.2 Mean proportions correct (± *SE*) for the similar, dissimilar, and CV_ lists of nonwords for the three scoring procedures (Experiment 2).

5.4.3. Discussion

When measured using the order accuracy criterion, order memory was better for the dissimilar as compared to similar lists, which was better than when the nonword lists shared the CV_ component. These findings replicate the results observed from study one into the effect that similarity has on order memory for words. Further, these findings are also consistent with the predictions generated from non-linguistic as well as psycholinguistic models of STM. For instance, non-linguistic models of STM that are based on the distinctiveness assumption (Nairne, 1988, 1990a), argue that as similarity increases order memory should decrease. Hartley and Houghton's (1996) linguistically constrained model of STM also predicts an order memory impairment for CV_ lists of nonwords, in that although phonemic similarity should have a detrimental effect on order memory (i.e., dissimilar as compared to phonemically similar lists), when the overlapping phoneme is the vowel (in this case, the CV_ condition) order memory should be further impaired.

When performance was measured for the recall of item information, more items were recalled when the nonword lists were phonemically similar as compared to CV_ lists, which was better than when the nonwords were phonemically dissimilar. According to Nairne (2002), an item recall advantage should be observed whenever the memory search set can be limited to a smaller number of items (in this case, CV_ lists). Hence, the finding of an item recall advantage for phonemically similar as compared to CV_ lists of nonwords is problematic for current versions of the feature model (Nairne, 2002). In contrast, the Hartley and Houghton (1996) model suggests that syllable representations do not receive the same level of support when list items share the CV_ as compared to _VC component. As a consequence, the item recall advantage observed for rhyming lists of nonwords should be absent (or at least minimal) when list items share the CV_ component. In addition, the rhyme ending is shared by some words in the similar condition in this experiment, so according to Hartley and Houghton's (1996) model, these items should benefit from the reinforcement of the relevant rhyme unit.

5.5. Experiment 3

Experiment 3 further examined the explanations that extant STM models generate for the effect that similarity has on order and item memory. Experiment 3 consisted of nonword lists that shared the C_C component. In terms of the effect that similarity has on order memory, non-linguistic STM models (Nairne, 1988, 1990a) predict an identical pattern of results to those observed across Experiments 1 and 2. Hence, these types of models predict that as similarity increases (in this case, lists of C_C nonwords) order memory should decrease. In contrast, the linguistically constrained model of STM developed by Hartley and Houghton (1996) suggests that, as compared to the phonemically dissimilar condition, any form of similarity should decrease order memory (i.e., C_C lists), however, when the overlapping phoneme is the vowel, order memory should be further impaired. Since all of the items in the phonemically similar condition in this experiment share the same vowel, the relatively greater weighting of the vowel as compared to consonants may actually produce poorer order memory for the phonemically similar in comparison to C_C lists.

At the item recall level, newer versions of the feature model (Nairne, 2002; see also Nairne & Kelley, 1999; Neath, 1999; Tehan & Fallon, 1999), suggest that when list items share features that can be used to limit the size of the search set (in this case, the C_C component), an item recall advantage should be observed. In contrast, the Hartley and Houghton (1996) model suggests that, as compared to when the stimulus lists rhyme, the syllable representations for C_C lists do not receive the same degree of support, since a separate rhyme unit is activated for each item. Therefore, the item recall advantage observed for rhyming lists of nonwords (i.e., Experiment 1) should be absent (or at least minimal) when the nonword lists share the C_C component.

5.5.1. Method

5.5.1.1. Participants

Twenty-four undergraduate psychology students from the University of Wollongong participant pool (2 males and 22 females), with an age range from 18 to 34 years (*M* = 22.63), participated in compliance with a course requirement. The same inclusion criteria were placed on the selection of participants as for Experiment 1.

5.5.1.2. Stimuli

The stimuli comprised 150 nonwords with a CVC structure (refer Appendix C – Table C3). The stimuli were used to create 30 same-consonant, 30 phonemically similar, and 30 phonemically dissimilar five-item lists. The same constraints were placed on the construction of the stimulus lists, as for Experiment 1, with two minor modifications. The first modification was that for the C_C lists all of the stimuli in a particular list shared the C_C component (e.g., *Bech, Barch, Borch, Baich,* and *Biech*). Also, for the phonemically similar lists, no item in a list shared both consonants (e.g., *Bech, Besh, Bedge, Teg,* and *Cheg*). Therefore, all of the items in a similar list shared the same vowel.

5.5.1.3. Procedure

The same testing procedure was used in the current experiment as in Experiment 1. The criterion for satisfactory audibility of the nonwords was met with participants correctly repeating 99.5% of the nonwords presented. Inter-rater reliability scores for the C_C (correct-in-position recall, 97%; item recall, 98%), similar (correct-in-position recall, 97%; item recall, 96%) and dissimilar (correct-in-position recall, 98%; item recall, 99%) conditions were obtained.

5.5.2. Results

The correct-in-position, item recall, and order accuracy measures were analysed using three separate repeated measures ANOVAs. Figure 5.3 summarises performance for the dissimilar, similar, and C_C conditions, collapsed across serial position for each of the three performance measures.

The correct-in-position analysis revealed a main effect of phonological similarity, $F(2,46) = 74.270$, $MSE = 136.051$, $p < .001$. Post hoc analyses revealed that C_C lists were recalled more accurately than either dissimilar, $t(23) = 11.193$, $p < .0167$, or similar lists, $t(23) = 8.157$, $p < .0167$, which did not differ, $t(23) = 2.372$, *ns.* The item recall analysis revealed a main effect of phonological similarity, $F(2,46) = 179.991$, $MSE = 98.403$, $p < .001$. Post hoc analyses revealed that item recall was higher for C_C as compared to similar lists, $t(23) = 11.084$, $p < .0167$, which was higher than when the nonwords were phonemically dissimilar, $t(23) = 8.275$, $p < .0167$. The order accuracy analysis revealed a main effect of phonological similarity, *Greenhouse-Geisser* $(1.553,35.729) = 55.875$, $MSE = .0085$, $p < .001$. Post hoc analyses revealed that dissimilar lists were recalled more accurately than C_C lists, $t(23) = 8.381$, $p < .0167$, which were more accurately recalled than similar lists, $t(23) = 3.343$, $p < .0167$.

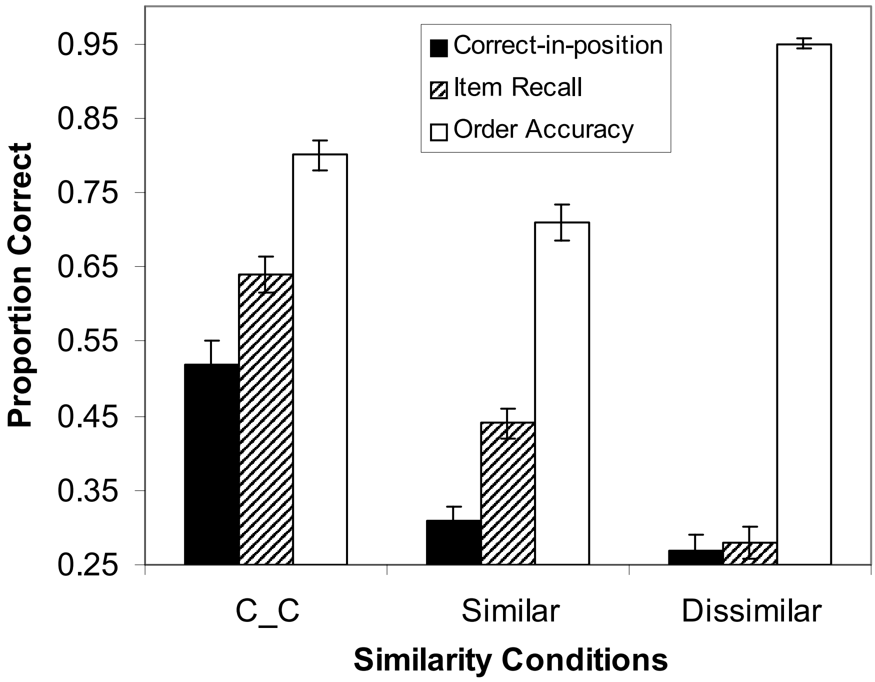

Figure 5.3 Mean proportions correct (± *SE*) for the similar, dissimilar and C_C lists of nonwords for the three scoring procedures (Experiment 3).

5.5.3. Discussion

Consistent with the predictions generated from the feature model (Nairne, 2002), when performance was measured for the recall of item information, more items were recalled in the C_C as compared to the phonemically similar condition, which produced more items than when the nonwords were phonemically dissimilar. In contrast, this finding of an item recall advantage for lists of C_C nonwords is problematic for Hartley and Houghton's (1996) linguistically constrained model of STM. According to this model, when list items share the C_C component, the syllable representations held in STM should not receive the same level of support in comparison to when the lists share a rhyme ending. As such, the recall of item information should not benefit to the same extent.

However, the measure of performance obtained using the item recall criterion is not independent from the order constraints placed on performance during serial recall

(Wickelgren, 1965d). For instance, in the current study, participants were told to recall the list items in the order in which they were presented. Therefore, in comparison to performance on tasks such as free recall (Bjork & Whitten, 1974; Glanzer & Cunitz, 1966), the measure of item recall ability obtained in the current study may have been influenced by the order constraints placed on performance during recall. Further, the current study was designed to examine performance across the three conditions within each experiment. As such, constraints that have been found to influence STM performance, such as sub-syllabic frequency (Nimmo & Roodenrys, 2002), item length (Baddeley et al., 1975), complexity (Service, 1998) and the ease of the articulatory transitions between list items (Murray & Jones, 2002) were not controlled across experimental sets.

One way to examine whether differences in the stimulus sets used across experiments differentially influenced the recall of item information is to compare performance across the three dissimilar conditions. A one-way ANOVA revealed a difference in item recall levels for the dissimilar conditions, $F(2, 71) = 3.458$, $MSE = 200.670$, $p < .05$. Although post hoc independent samples *t*-tests revealed no significant differences between any of the dissimilar conditions (Experiment 1 vs. 2, $t(46) = 1.310$, *ns*; Experiment 1 vs. 3, $t(46) = 2.472$, *ns*; Experiment 2 vs. 3, $t(46) = 1.408$, *ns*), the significance of the overall test suggests that there are some differences between the experimental sets. As a result, it seems unwise to further examine the effect that phonemic similarity (when similarity is held constant - _VC, CV_ and C_C lists) has on the recall of item information.

In terms of the effect that similarity has on order recall (using the order accuracy measure), the current study found that order memory was better for dissimilar as compared to C_C lists, which was better than when the stimuli were phonemically similar. This is an important finding in that non-linguistic STM models (Nairne, 1988, 1990a) assume that as similarity increases order memory should decrease. However, in the current experiment, C_C lists shared a greater amount of phonemic overlap than did the phonemically similar lists of nonwords. In contrast, Hartley and Houghton's (1996) linguistically constrained model of STM suggests that, although any form of similarity should impair order memory, when the overlapping phoneme is the vowel (the phonemically similar condition in the current experiment), order memory should be further impaired.

Unlike the measure of order performance obtained using the correct-in-position criterion, the order accuracy measure takes into account individual differences in the recall of item information. This is notable in that the item recall levels for the dissimilar conditions observed in the current study were found to vary across experiments. However, when the scores obtained using the order accuracy criterion were compared across the dissimilar conditions, similar levels of performance were found (Experiment 1, .93; Experiment 2, .93; Experiment 3, .95). This was confirmed by performing a one-way ANOVA on the order accuracy data which revealed no differences across the phonemically dissimilar conditions, $F(2, 71) = .905$, $MSE = .004$, $p > .05$. This finding suggests that meaningful comparisons can be made across experiments when performance is measured using the order accuracy criterion.

To examine the effect that similarity has on order memory in greater detail, the data for the similar conditions, where the position of the overlapping phonemes was varied, yet similarity remained constant (i.e., _VC, Experiment 1; CV_, Experiment 2; C_C, Experiment 3), were subjected to a one-way ANOVA. This analysis revealed a significant difference in performance across experiments, $F(2, 71) = 19.244$, $MSE = .011$, $p < .05$. Post hoc independent samples *t*-tests revealed that order memory was better for C_C as compared to either rhyming, $t(46) = 3.854$, $p < .0167$, or CV_ lists, $t(46) = 6.064$, $p < .0167$, which did not differ, $t(46) = 2.395$, *ns*. Hence, the predictions generated from non-linguistic models of STM (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a, 2002) that predict similar levels of performance across lists that are phonemically equivalent (i.e., _VC, CV_ or C_C lists), are clearly inconsistent with the current research findings. In contrast, the Hartley and Houghton (1996) model suggests that although any form of similarity should have a detrimental effect on order memory, when the overlapping phoneme is a vowel, order memory should be further impaired. Hence, these findings lend strong support to Hartley and Houghton’s (1996) linguistically constrained model of STM for nonwords.

5.6. General Discussion

The aim of the current study was to test the predictions that STM models generate for the effect that phonemic similarity has on the recall of both item information, and memory for an item's position in a list when the experimental stimuli are nonwords. In terms of the recall of item information, neither the feature model (Nairne, 2002), nor a model of STM that places linguistic constraints on STM performance (Hartley & Houghton, 1996), could adequately deal with the current research findings. However, as suggested previously the measure of item recall obtained in the current study may have been influenced by the order constraints that are placed on performance when the task is serial recall. Although beyond the scope of the current study, to draw sound conclusions with respect to the effect that similarity has on item recall, future research should aim to employ a purer item recall measure: a measure that is unconstrained by order requirements during recall.

Consistent with a recent study conducted by Fallon et al. (in press), when performance was scored using the order accuracy criterion, order memory was better for phonemically dissimilar lists of nonwords, in comparison to any of the similar conditions. The finding of an order memory impairment for phonemically similar lists of items mirrors the findings observed from studies that have employed words (see Study 1). Of greater interest was the finding that order memory (as measured using the order accuracy criterion) was better for C_C (Experiment 3), as compared to either rhyming (Experiment 1) or CV_ (Experiment 2) lists of nonwords. This is important in that for these three phonemically similar conditions, the amount of phonemic overlap shared between lists was held constant across conditions. Currently, non-linguistic models of STM (Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a, 2002) assume that the locus of the similarity effect for order memory derives from the amount of phonemic overlap that is shared between list items. Hence, these types of models argue that, if similarity is held constant across lists, order memory should be impaired to the same extent, regardless of which of the phonemes are shared between list items (i.e., CV_, C_C or _VC components). Therefore, in terms of the effect that similarity has on order memory, the findings from the current study are clearly inconsistent with non-linguistic models of STM that are based on the distinctiveness assumption (Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a, 2000). Rather, the results are consistent with the predictions generated by Hartley and Houghton's (1996) linguistically constrained model of STM. According to this model, sharing any form of similarity with other list items should make it harder to recall the items in the correct order (i.e., phonemically dissimilar as compared to any of the similar conditions in the current study), however, when the overlapping phoneme is the vowel (i.e. CV_ and _VC, as compared to C_C lists), a further impairment in order memory should be observed. Hence, the findings from the current study are more consistent with the idea that when the experimental lists are phonemically similar, order memory is influenced by both the amount of phonemic overlap between list items and whether the overlapping phoneme is the vowel.

In summary, Hartley and Houghton's (1996) linguistically constrained model of STM is currently the only STM model that can adequately explain the current research findings. However, this model is limited in that it was developed to explain findings from a small body of research on STM for nonwords. Recently, Burgess and Hitch (1999) have suggested that the Hartley and Houghton (1996) syllable template mechanism could be incorporated into their latest connectionist model of the phonological loop (Although this would necessitate the inclusion of item representations at the sub-syllabic level, in which the parts of a syllable [i.e., the onset and rhyme nodes] could be temporally ordered). The amalgamation of these two models would provide researchers with an STM model that could not only serially order list items and explain a wide variety of other recall effects (Burgess & Hitch, 1999), but also provide a mechanism that was capable of linguistically constraining STM performance at the sub-syllabic level. As such, a more fruitful avenue in which to guide future STM models may be to modify the Burgess and Hitch (1999) model to include Hartley and Houghton's (1996) syllable template mechanism.

To reiterate, the major implications of the current findings are as follows. Firstly, it is no longer sufficient to use a 'phonemic overlap' or distinctiveness argument to account for the effect that phonemic similarity has on memory for an item's position in a list. Rather, the findings suggest that it is the consistency and the phonemic make up of the list items which influences STM performance. Secondly, STM performance is influenced by linguistic constraints, such as syllable structure and sonority that are assumed to operate at the sub-syllabic level. Hence, the current research points to a desperate need for STM modellers to incorporate the mechanisms necessary to deal with the psycholinguistic rules that constrain speech production at the sub-syllabic as compared to lexical level.

PRELUDE TO CHAPTER 6

One broad question that research with nonwords attempts to address is whether words and nonwords are processed in the same way. Currently, this is a contentious issue with some researchers suggesting that different STM processes are involved in word as compared to nonword recall (Lian et al., 2001; Saint-Aubin & Poirier, 2000), whereas others (e.g., Fallon et al., in press) are less convinced. Evidence in support of the idea that words and nonwords are processed differently stems from the finding that when performance is measured using the correct-in-position criterion, phonological similarity has a detrimental effect on order memory when the experimental stimuli are words (Baddeley, 1966). However, when the stimuli are nonwords, phonemic similarity has been found to facilitate order memory (Fallon et al., in press; see also Study 2).

Although these types of findings may be used as evidence to support the idea that different processes are involved in word and nonword recall, as Wickelgren (1965d) suggests, the measure of order memory commonly used by researchers (i.e., correct-in-position) is not independent of an individual's item recall ability. This is compounded by the fact that overt speech production constraints on performance influence the number of items an individual is able to recall (Snowling, 1989; Wells, 1995). According to Gathercole et al. (1999), this appears to be especially true when the experimental stimuli are nonwords.

The influence that item recall ability has on the results obtained when the traditional measure (i.e., correct-in-position) is used to assess order memory can be demonstrated by examining the findings observed when order memory is measured using the order accuracy criterion (i.e., the measure of order memory that controls for the effect that individual differences in item recall has on the number of order errors). When the effect that phonological similarity has on order memory is scored using the order accuracy criterion, the standard order memory impairment for phonemically similar as compared to dissimilar lists of items is observed, regardless of whether the stimuli are words (Fallon et al., 1999; see also Study 1) or nonwords (Fallon et al., in press; see also Study 2). Hence, to investigate whether the same (Fallon et al., in press) or different (: #150)(Lian et al., 2001; Saint-Aubin & Poirier, 2000) processes are involved in word as compared to nonword recall, findings from serial recall tasks should be contrasted with the results observed from a task that does not require participants to verbally recall the presented list items.

There were three main aims for conducting the current study. The first aim was to investigate whether the detrimental effect that phonemic similarity has on order memory persists, once the demands that overt speech production have on STM performance are removed. Secondly, to examine whether the effect that phonemic similarity has on order memory is similar, regardless of whether the experimental stimuli are words or nonwords. The final aim of the current study, as with Studies one and two, was to critically examine the predictions that non-linguistic (e.g., Nairne, 1988, 1990a, 2002) and psycholinguistic (e.g., Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) models of STM generate for the effect that similarity has on order memory.

6. THE PHONOLOGICAL SIMILARITY EFFECT IN SERIAL RECOGNITION*

By

Lisa M. Nimmo and Steven Roodenrys

University of Wollongong
Wollongong, Australia

Address for correspondence:
Lisa Nimmo, Department of Psychology, University of Wollongong
Northfields Ave, Wollongong, 2522, Australia
Phone: (+61) 242 214 454, Fax: (+61) 242 214 163
Email: lmn02@uow.edu.au

Running head: Phonological similarity and serial recognition

*This paper is currently under review (Memory)

6.1. Abstract

The aim of the current research was to determine whether the phonological similarity effect is influenced by the consistency and position of phonemes within list items in a serial recognition task. Three experiments with consonant-vowel-consonant (CVC) items are reported, where the position of the phonemes that items shared were manipulated (i.e., list items shared the vowel and final consonant [_VC; Experiment 1], shared the initial consonant and vowel [CV_; Experiment 2], or the two consonants [C_C; Experiment 3]. The results show that regardless of whether the stimuli are words or nonwords, the influence of sub-syllabic mechanisms on short-term memory (STM) performance is independent of speech production processes. The results suggest that the same mechanisms subserve the recall of words and nonwords in STM and that overt speech production processes influence the findings obtained from studies that require the verbal recall of nonword lists. Implications for current STM models are discussed.

6.2. Introduction

One of the most widely used measures of short-term memory (STM) is the immediate serial recall task (Gathercole et al., 2001). The serial recall task requires participants to verbally recall lists of items, in the order in which they were presented. An extremely robust effect in this task is that recall of lists of words is much better than recall of lists of nonsense syllables (e.g., Hulme et al., 1991; Hulme et al., 1995) Termed *the lexicality effect*, this performance advantage has been used as evidence to support the notion that unlike nonwords, words undergo a redintegration process (Brown & Hulme, 1995) where, prior to output, pre-existing long-term memory (LTM) representations aid in the recall of incomplete traces held in STM (Schweickert, 1993). Although this view has recently been tempered with the suggestion that nonwords that are high in word-likeness may also undergo a redintegration process (see Saint-Aubin and Poirier, 2000).

Another factor that has been found to influence STM performance is the phonological similarity of the experimental list items. The phonological similarity effect is the finding that STM performance is worse if the words in a list sound similar to each other (e.g., Baddeley et al., 1984; Coltheart, 1993; Conrad & Hull, 1964; Fallon et al., 1999; Gathercole et al., 1982; Wickelgren, 1965d). More recent research on the effect that phonemic similarity has on STM performance has extended this detrimental finding to situations where the experimental stimuli are nonwords (e.g., Fallon et al., in press; Gathercole et al., 2001).

Although the majority of STM models include mechanisms to account for the effect that both phonemic similarity and lexicality have on STM performance (e.g., Brown et al., 2000; Burgess & Hitch, 1992; 1999; Gupta & MacWhinney, 1997; Henson, 1998; Nairne, 1988, 1990a; Tehan & Fallon, 1999), a point of difference between these models is in the locus at which they assume these effects occur.

6.2.1. Current short-term memory models

A core assumption proposed to account for the lexicality effect is that the redintegration process operates at the lexical level and, according to some models, the effect that phonemic similarity has on order memory is a consequence of this process (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999). By this view, pre-existing LTM representations can be used to aid in the recall of degraded STM traces. Words are assumed to benefit from these LTM representations during recall, whereas nonwords are not. As such, researchers have suggested that phonemic similarity should not influence performance when the experimental stimuli are nonwords (e.g., Brown & Hulme, 1995). However, some research findings are inconsistent with this view. For instance, Besner & Davelaar (1982; see also Gathercole et al., 2001) found an order memory impairment for phonemically similar as compared to dissimilar lists of nonwords. This finding is problematic for STM models that assume that the phonological similarity effect arises purely from pre-existing lexical representations competing in redintegration (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999).

However, a number of STM models do provide an explanation for the effect of phonemic similarity on nonword recall by assuming that similar processes operate for word and nonword recall (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996; Nairne, 1988, 1990a, 2002). This idea is theoretically consistent with earlier conceptualisations of STM processes. For instance, the evolutionary utility of the phonological loop component of the working memory model stems from the belief that this component is responsible for vocabulary acquisition (Baddeley et al., 1998). According to this view, the difference between word and nonword recall is one of degree, with item familiarity being the influential variable. This idea is also consistent with research that suggests that the more quickly participants are able to associate a real word with a nonword (Lian et al., 2001), or the more word-like a nonword is rated (Gathercole et al., 1991b; Gathercole, 1995; Gathercole & Martin, 1996; Metsala, 1999), the more accurately the nonword is recalled.

There are two types of STM models that provide explanations for the effect of phonemic similarity on nonword recall: non-linguistic and psycholinguistic models of STM. Non-linguistic STM models were designed to examine general memory principles, and as such place no emphasis on stimulus type (e.g., words, nonwords, pictures or spatial location). An example of this type of model is Nairne’s (1988, 1990a, 2002) feature model of immediate memory. This can be contrasted with psycholinguistic models of STM that were specifically designed to examine language-based constraints on STM performance (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996).

Regardless of stimulus type (i.e., words or nonwords), both classes of model predict an order memory impairment for phonemically similar as compared to dissimilar lists of items. However, as described below, the mechanisms that each class of model proposes to account for this effect differ.

6.2.1.1. Non-linguistic short-term memory models

Like other STM models (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999), the feature model (Nairne, 1988, 1990a, 2002) is based on the ratio rule or distinctiveness assumption. This is the idea that the likelihood of recalling a target item is relative to the similarity of the target item to all of the other items that were presented in the same list (Luce, 1959; Nosofsky, 1986). Accordingly, STM models that are based on the distinctiveness assumption argue that as phonemic similarity between list items increases order memory should decrease. Further, newer versions of the feature model (e.g., Nairne, 2002; Neath, 1999) assume that the long-term representations that aid in the redintegration process are formed during the experimental session. In other words, these long-term representations are context specific. As such, the feature model suggests that the effect that phonemic similarity has on order memory should be unaffected by the lexicality of the stimuli.[13]

6.2.1.2. Psycholinguistic models of short-term memory

STM models that can be classified as psycholinguistic are based on the assumption that unique demands are placed on memory when recalling verbal as compared to other types of material (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996; Martin et al., 1999). Unlike non-linguistic STM models, psycholinguistic models of STM assume that sub-syllabic linguistic processes are

[13] The feature model treats word and nonword recall in the same way. Hence, it is unclear how current versions of the feature model could explain the lexicality effect.

responsible for the effect that phonemic similarity has on order memory. One such model has been developed by Hartley and Houghton (1996).

One linguistic principle that Hartley and Houghton (1996) have incorporated into their model is the idea of sonority. Sonority refers to the energy of a speech trace, and the sonority principle, to the idea that the energy of a speech trace for syllables increases to a peak at the vowel and then decreases (Treiman, 1984). Accordingly, the speech trace for vowels as compared to consonants will be stronger because the trace is both longer in duration and more acoustically intense than it is for consonants (Hartley & Houghton, 1996; see Service, Maury & Luotoniemi, in press, for a similar argument). To model sonority, Hartley and Houghton (1996) set the activation level of nodes representing vowels higher than the nodes that represent consonants. Therefore, according to this model, any form of similarity should impair order memory but the greatest impairment will be seen when the vowel is shared, as this is the most strongly represented phoneme in a speech trace. However, this model is limited in that it was designed to explain research findings based on nonword experiments.

In contrast, Gupta and MacWhinney (1997) have developed an STM model that deals with the recall of both words and nonwords. This model is based on a combination of the Burgess and Hitch (1992) connectionist model, and Hartley and Houghton's (1996) linguistically constrained model of STM. As such, this hybrid model provides a mechanism that can serially order items (i.e., Burgess & Hitch, 1992) and uses sub-syllabic linguistic processes to constrain STM performance (e.g., Hartley & Houghton, 1996). According to this hybrid model of vocabulary acquisition, the speech trace held in the phonological chunk layer is more stable for words as compared to nonwords. This is due to the added support that words receive from both a semantic layer and the phonological store, which are both assumed to contain pre-existing LTM representations. However, in contrast to STM models that are based on the distinctiveness assumption, this hybrid model suggests that STM performance is also influenced by linguistic mechanisms that operate at the sub-syllabic level. Hence, according to the Gupta and MacWhinney (1997) model, although overall performance should be higher for words as compared to nonwords, phonological similarity should influence order memory in a similar way, regardless of stimulus type. That is, any form of phonemic similarity between list items should have a detrimental effect on order memory. However, when the overlapping phoneme is the vowel, a further order memory impairment should be observed as vowels are more strongly represented than consonants.

One difficulty in evaluating STM model accounts of the phonological similarity effect is the inconsistency of findings in the research literature. For example, although the standard detrimental effect that similarity has on order memory is observed when performance is measured using the correct-in-position criterion (i.e., scored as correct if a participant recalls the correct item in the correct position) for words (e.g., Baddeley, 1966; see study 1), when the stimuli are nonwords, phonemic similarity has been found to facilitate order memory (e.g., Fallon et al., in press, see study 2). These findings suggest either that different STM processes are involved when recalling lists of words as compared to nonwords (Brown & Hulme, 1995), or that other factors inherent in verbal recall tasks, such as speech production processes, influence the results obtained when the experimental stimuli are nonwords.

6.2.2. Speech output processes and serial recall

Although not a new concept, one factor that has been found to influence serial recall performance is verbal speech production constraints (Gathercole et al., 1999; Snowling & Hulme, 1989; Wells, 1995). A related issue is the idea that when participants are required to perform the serial recall task, measures of order (correct-in-position) and item recall (i.e., scored as correct if a participant recalls a presented item, regardless of the position in which it was recalled) used to assess STM performance are not independent (Wickelgren, 1965d). This is important in that the number of items an individual is able to recall may be influenced by how easy (e.g., *wut, zut, vut, yut, chut*) or difficult (e.g., *zut, yied, hig, chone, wayb*) the experimental lists are to pronounce. Consistent with this idea, recent research suggests that the ease of the articulatory transitions between list items influences recall performance (Murray & Jones, 2002; see also Service & Maury, 2003). Further, when the effect that phonemic similarity has on order memory was measured using the order accuracy criterion - a measure of order memory that controls for the effect that individual differences in item recall has on the number of order errors - we found the standard order memory impairment for phonemically similar as compared to dissimilar lists of items, regardless of whether the stimuli were words (Study 1) or nonwords (Study 2).

To reiterate, when order memory is measured using the correct-in-position criterion and the experimental stimuli are words, phonemic similarity impairs order memory (Baddeley, 1966), whereas for nonwords, a facilitative effect has been found (Fallon et al., in press). However, when measured using the order accuracy criterion, we found that the effect that phonemic similarity has on order memory was identical for words and nonwords (Study 1 & 2 respectively). These findings suggest that the inconsistencies observed in the research literature may be due to the difficulties associated with articulating nonwords, rather than to differences in the way in which words and nonwords are processed. Hence, to investigate whether the same processes are involved when recalling words as compared to nonwords, findings from serial recall tasks should be contrasted with the results observed from a task that is unconstrained by the processes involved in overt speech production.

One task that does not require that participants verbally recall presented list items, yet is similar to the standard serial recall task, is the serial recognition task (Campbell & Butterworth, 1985; Martin & Breedin, 1992; Martin et al., 1999). This task involves presenting a list of items to participants and then re-presenting either the same items in the same order, or the same items in a different order (Martin & Breedin, 1992). A participant's task is to say whether the items that were re-presented were in the 'same' or a 'different' order (e.g., *log, bog, hog* and then *bog, log, hog*). As this task is resistant to speech production errors, Gathercole et al. (2001) have recently suggested that the serial recognition task may be a better STM measure than the standard serial recall task.

To date, only two studies have examined the phonological similarity effect with the serial recognition task. In one study, Gathercole et al. (2001) found the standard phonological similarity effect (i.e., better recognition of dissimilar than similar lists) regardless of whether the stimuli were words or nonwords (Experiments 3B & 4B). However, although Lian et al. (2001) found an identical pattern of results for word lists, when nonwords were rated as being low in associative value (i.e., low in wordlikeness), they found no differences in recognition performance between phonemically dissimilar as compared to similar lists of nonwords. However, Lian et al. (2001) did not operationally define what they meant by phonemic similarity. This is important in that the findings from Studies one and two suggest that, when similarity is held constant (as measured by the degree of shared consonant and vowel information) and performance is measured using the order accuracy criterion, sharing a vowel impairs order memory to a greater extent than when the shared phonemes are consonants, regardless of whether the stimuli are words or nonwords.

6.2.3. The current experiments

Three experiments were designed to assess whether the influence that phonemic similarity has on order memory is independent of an individual's overt speech production abilities. Given that previous research (e.g., Studies 1 & 2) suggests that the phonemic similarity effect is influenced by sub-syllabic linguistic processes, similarity was defined in a number of different ways. For instance, Experiment 1 consisted of lists of words and nonwords that rhymed, whereas the important condition for Experiment 2 consisted of words and nonwords that shared a common consonant-vowel (CV_) component. Thus, changing the position of the overlapping phonemes across experiments, yet keeping the amount of phonemic overlap constant. Finally, the important condition in Experiment 3 consisted of words and nonwords that shared common initial and final consonants (C_C). To make the comparison easier between the current research findings and previous research that has used the serial recall task (Studies 1 & 2), participants were auditorily presented with lists of items at a fixed length.

There were three main aims for conducting the current research. The first aim was to assess whether the effect that phonemic similarity has on order memory is found when the task does not require overt speech production. The second aim of the current research was to compare the effect of phonemic similarity on order memory for words and nonwords. If the pattern of effects varies with stimulus type it would suggest that different STM processes are operating for word as compared to nonword recall. Alternatively, a similar pattern of results would suggest that similar processes are involved in word and nonword recall, and that the differential results observed in the research literature (see Studies 1 & 2) using the correct-in-position criterion, are due to differences in overt speech production demands between words and nonwords.

The final aim of the current research was to distinguish between the predictions that non-linguistic and psycholinguistic models of STM generate for the effect of phonemic similarity on order memory. For instance, STM models that do not incorporate sub-syllabic linguistic constraints on STM performance (e.g., Nairne, 1988, 1990a) suggest that as similarity increases order memory should decrease. In contrast, psycholinguistic models of STM (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) predict that any form of phonemic similarity should impair order memory but the greatest impairment will be seen when the overlapping phoneme is the vowel.

6.3. Experiment 1

6.3.1. Method

6.3.1.1. Participants

Seventy-two undergraduate psychology students from the University of Wollongong participant pool (18 males and 54 females), with an age range from 17 to 37 years (*M* = 20), participated in compliance with a course requirement. Only native Australian English speakers who indicated having no prior problems with hearing participated in the study.

6.3.1.2. Stimuli

Two stimulus sets that comprised either words or nonwords were constructed.

Word pool. The word pool comprised 180 words with a consonant-vowel-consonant (CVC) structure (refer Appendix A – Table A1). The stimuli were used to create 30 rhyming, 30 phonemically similar and 30 phonemically dissimilar six-word lists. Thus, each word was sampled three times, such that each word appeared in one rhyming, one phonemically similar and one phonemically dissimilar list. For the rhyming condition, all of the stimuli in a particular list shared the _VC component (e.g., *Name, Came, Maim, Lame, Dame,* and *Shame*). For the similar condition, two constraints were placed on list construction. The first constraint was that no items in a list shared the _VC component. Also, each stimulus in each list shared two phonemes with at least one other stimulus in the same list (e.g., *Name, Knock, Need, Knees,* *Gnome,* and *Not*). Hence, all of the items in the similar condition shared the same initial consonant. For the dissimilar condition, each stimulus in each list did not share any phonemes with any other stimulus in that list (e.g., *Name, Cot, Soul, Died, Pig,* and *Hub*).

Nonword pool. The nonword pool comprised 180 nonwords with a CVC structure (refer Appendix E – Table E1). The same constraints were placed on list construction for the nonwords as is outlined when the experimental stimuli were words.

Using an Arista Cardioid dynamic microphone (Model number DM-904D), the stimuli were recorded using a Sony Minidisc Deck (Model Number: MDS-JE640) in a sound attenuated booth by a female speaker with an Australian English accent. Each stimulus was transferred digitally onto a Macintosh computer and normalised to control for possible amplitude effects on recognition performance.

6.3.1.3. Procedure

Thirty participants were presented with stimulus lists that were composed of words. The remaining 42 participants were presented with stimulus lists that were composed of nonwords. Thus, lexicality was a between subjects factor. Regardless of the lexical status of the items, across all conditions, two practice lists were given to each participant prior to the presentation of the first experimental list. The lists were presented in three blocks of thirty trials. The order of the blocks within the experimental session was counterbalanced across participants. The order of the trials in each block and the order in which the items occurred in each list were randomised for all participants.

Each participant was auditorily presented with either six words or nonwords (target list) at a rate of one item per second. The same six items were then re-presented to each participant (comparison list). Stimulus presentation rate was controlled using Hypercard (version 2.4.1). One second after the presentation of the last item in a list, participants heard a 200 millisecond, 500 Hz tone which was used as a signal for participants to respond. The participant's task was to say whether the items in the comparison list were in the same (15 lists) or a different order (15 lists) to the target list. Hence, the only difference between the two list presentations was whether the items in the comparison list were presented in exactly the same order as the target list or whether two of the list items had been transposed. The number of transpositions was counterbalanced across each condition. Thus, of the 15 comparison lists where two of the items had been transposed, 3 transpositions occurred at each position (i.e., position 1 with 2, 2 with 3, 3 with 4, 4 with 5 and 5 with 6). The participant's response was scored immediately after each comparison list was presented. The time taken for each participant to complete all three conditions was approximately 40 minutes.

6.3.2. Results

The mean number of correct responses was collated for words and nonwords when the stimuli were phonemically dissimilar, similar or rhymed (refer Table 6.1).

Table 6.1 Mean Proportions (with Standard Deviations) of Lists Correctly Recognized in Experiment 1 as a Function of Lexicality and List Type.

	Rhyming		*Phonemically Similar*		*Phonemically Dissimilar*	
Lexicality	*M*	*SD*	*M*	*SD*	*M*	*SD*
Words	0.62	0.09	0.63	0.07	0.70	0.10
Nonwords	0.60	0.10	0.63	0.10	0.68	0.12
Mean	0.61		0.63		0.69	

To control for potential shifts in bias, or levels of attention for each participant, hits and false alarm rates were collated and used to calculate a *d* prime (*d'*) value. The *d'* data were analysed using a 2 (lexicality) x 3 (phonological similarity) mixed design analysis of variance (ANOVA).[14] Figure 6.1 summarises the resulting *d'* values for words and nonwords when the stimuli were phonemically dissimilar, similar or rhymed. Unless otherwise specified, alpha was set at .05 (2-tailed), and an identical pattern of

[14] Refer Appendix F for the ANOVA tables for the major statistical analyses performed for study three.

results was found regardless of whether the analyses were performed on the number of correct trials or the mean *d'* values obtained.

The *d'* analysis revealed a main effect of phonological similarity, $F(2,140) = 14.571$, $MSE = .349$, $p < .01$, but no effect of lexicality, $F(1,70) = .170$, $MSE = .633$, *ns* refer Figure 6.1). Phonological similarity was not found to interact with lexicality, $F(2,140) = .164$, *ns.* Post-hoc paired samples *t*-tests were used to analyse the main effect of phonological similarity. Alpha was set at 0.0167 (2-tailed) to control for the increased probability of committing a Type I error as a function of the number of comparisons performed, thus keeping the family-wise error rate at .05. The analyses revealed that order memory was better for phonemically dissimilar as compared to either of the similar conditions (rhyming, $t(71) = 4.602$, $p <.0167$; phonemically similar, $t(71) = 4.246$, $p < .0167$), which did not differ, $t(71) = .843$, *ns.*

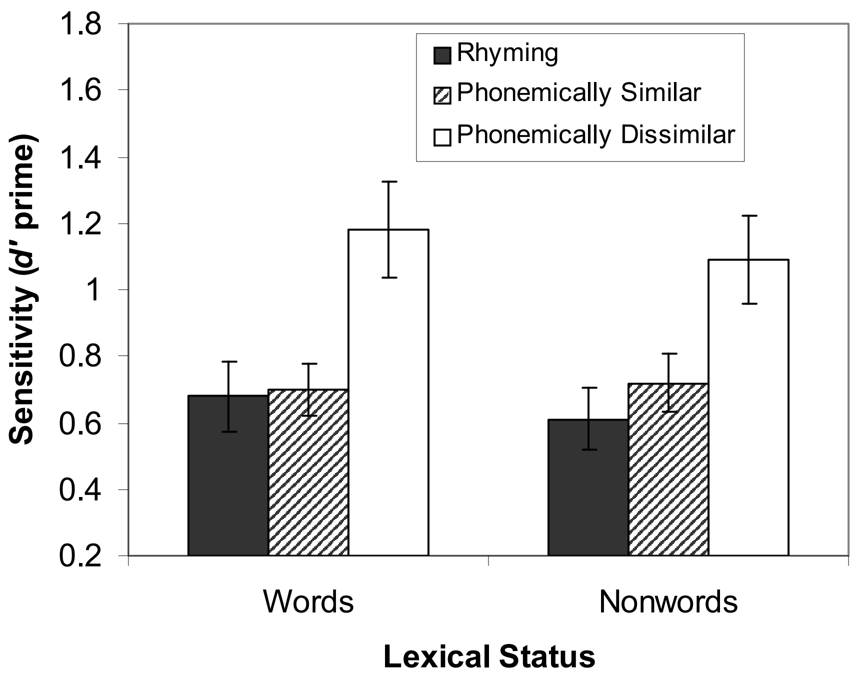

Figure 6.1 Means (± SE) for phonemically similar, phonemically dissimilar and rhyming lists as a function of lexicality for Experiment 1, using a measure of sensitivity (*d'* prime).

6.3.3. Discussion

Consistent with a recent study conducted by Gathercole et al. (2001), regardless of whether the stimuli were words or nonwords, phonemic similarity had a detrimental effect on order memory. In other words, order memory was better for the phonemically dissimilar as compared to either rhyming or phonemically similar lists of stimuli, which did not differ. Hence, the findings observed in the current study suggest that the effect of phonemic similarity on order memory remains, once the demands that overt speech production has on STM performance are removed.

Further, the findings from the current study are inconsistent with the explanations that non-linguistic (Nairne, 1988, 1990a) and psycholinguistic (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) models of STM generate to account for the effect of phonemic similarity on order memory. For instance, according to non-linguistic STM models(Nairne, 1988, 1990a), as phonemic similarity increases order memory should decrease. In contrast, the current study found no order memory differences between phonemically similar and rhyming lists of stimuli, despite the greater similarity in the rhyming lists. Further, psycholinguistic models of STM (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) suggest that any form of phonemic similarity should have a detrimental effect on order memory but the greatest impairment will be seen when the vowel is shared (i.e., in this case, rhyming lists of stimuli).

The findings from the current study are, however, consistent with those from a recent word study, in that we (see study 1; Experiment 1) found an identical pattern of results when performance on a serial recall task was measured using the correct-in-position criterion. In contrast, when the experimental stimuli were nonwords, we (Study 2; Experiment 1) found an order memory advantage for items that shared some form of phonemic similarity in comparison to when the experimental lists consisted of phonemically dissimilar nonwords. These earlier findings suggested that words and nonwords may be processed differently. Hence, the findings from the current study may help resolve the inconsistencies in the results observed between words and nonwords in serial recall.

Using serial recognition, which does not require overt speech production, the current study found an identical pattern of results, regardless of whether the stimuli were words or nonwords. These findings suggest that the differences observed in serial recall when the stimuli are words as compared to nonwords may not be a result of different STM processes. Rather, the results of the current experiments suggest that the same processes are involved in word and nonword recall and that it is the articulatory ability of participants that influences the results obtained for serial recall of nonwords (Gathercole et al., 1999).

Finally, there was no effect of lexicality on overall performance level. This finding of no effect of lexicality when the task is serial recognition is not entirely surprising. For instance, although Gathercole et al. (2001; Experiment 1) found an overall lexicality effect, no differences in serial recognition performance were observed when the lists contained fewer than six items. According to Gathercole et al. (2001), the effect that lexicality has on STM performance should be markedly reduced when the task is serial recognition as compared to serial recall. This is due to the fact that unlike serial recall, which requires that participants remember both item and order information, performance on the serial recognition task is less reliant on the retention of item information (Gathercole et al., 2001).

6.4. Experiment 2

A further two experiments will be discussed that were designed to investigate whether the effect of phonemic similarity on order memory is independent of overt speech production requirements. The second aim of the current experiments was to examine the explanations that non-linguistic (Nairne, 1988, 1990a) and psycholinguistic (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) models of STM generate for this effect.

Experiment 2 consisted of lists of words and nonwords that shared the CV_ component, whereas Experiment 3 consisted of words and nonwords that shared the C_C component. Hence, changing the position of the shared phonemes that list items overlap on, while keeping the amount of phonemic overlap (i.e., similarity) constant. If it is the case that the effect of similarity on order memory (as measured using the correct-in-position criterion) when the experimental stimuli are nonwords is influenced by the processes involved in speech production, then the results observed in the following experiments should replicate the correct-in-position results obtained in study one (Experiment 2). Hence, the expectation is that serial recognition performance should be better when the stimuli are phonemically dissimilar as compared to the phonemically similar lists, which in turn should be better than when the stimulus lists share the CV_ component.

Further, although non-linguistic (Nairne, 1988, 1990a) and psycholinguistic (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) models of STM differ with regards to the mechanisms each class of model assumes is responsible for the effect of phonemic similarity on order memory, the predictions they generate for the following experiment are indistinguishable. For instance, non-linguistic STM models predict that as similarity increases order memory should decrease. Psycholinguistic models of STM also predict an order memory impairment. According to these models, any form of phonemic similarity (i.e., in this case, the phonemically similar condition) should have a detrimental effect on order memory but the greatest impairment should be observed when the stimuli share the vowel (i.e., CV_ lists in the current experiment).

6.4.1. Method

6.4.1.1. Participants

Seventy-two undergraduate psychology students from the University of Wollongong participant pool (20 males and 52 females), with an age range from 18 to 46 years (*M* = 22), participated in compliance with a course requirement. The same inclusion criteria were placed on the selection of participants for the current experiment as in Experiment 1.

6.4.1.2. Stimuli

Two stimulus sets comprising of either words or nonwords were constructed.

Word pool. The word pool comprised 180 words with a CVC structure (refer Appendix A – Table A2). The stimuli were used to create 30 CV_, 30 phonemically similar and 30 phonemically dissimilar six-word lists. List construction was the same as in Experiment 1, except for a few minor modifications. The first modification was that for the CV_ condition, all of the words in a particular list shared the CV_ component (e.g., *Time, Ties, Tight, Type, Tide,* and *Tile*). Also, for the similar condition, no items in a list shared the CV_ component (e.g., *Time, Rum, Rhyme, Lime, Limb,* and *Dumb*). Hence, all of the items in a phonemically similar list shared the same final consonant.

Nonword pool. The nonword pool comprised 180 nonwords with a CVC structure (refer Appendix E – Table E2). The same constraints were placed on list construction for the nonwords as outline above.

6.4.1.3. Procedure

The same testing procedure was used in the current experiment, as in Experiment 1.

6.4.2. Results

The same analyses were performed on the data obtained in the current experiment as in Experiment 1. The mean proportions of correct responses were collated for words and nonwords when the stimuli were phonemically dissimilar, similar or shared the CV_ component (refer Table 6.2).

Table 6.2 Mean Proportions (with Standard Deviations) of Lists Correctly Recognized in Experiment 2 as a Function of Lexicality and List Type.

	CV_ lists		*Phonemically Similar*		*Phonemically Dissimilar*	
Lexicality	*M*	*SD*	*M*	*SD*	*M*	*SD*
Words	0.57	0.13	0.65	0.09	0.71	0.13
Nonwords	0.58	0.08	0.64	0.10	0.68	0.10
Mean	0.58		0.64		0.70	

Figure 6.2 summarises the mean *d'* values for words and nonwords when the stimuli were phonemically dissimilar, similar or shared the CV_ component. The *d'*

analysis revealed a main effect of phonological similarity, $F(2,140) = 30.317$, $MSE = .358$, $p < .001$, but no effect of lexicality, $F(1,70) = .001$, $MSE = .727$, *ns.* Phonological similarity was not found to interact with lexicality, $F(2,140) = .761$, *ns.* Post-hoc analyses on the main effect of similarity revealed that order memory was better for phonemically dissimilar as compared to the phonemically similar lists, $t(71) = 3.839$, $p < .0167$, which was better than when the experimental lists consisted of items that shared the CV_ component, $t(71) = 4.385$, $p < .0167$.

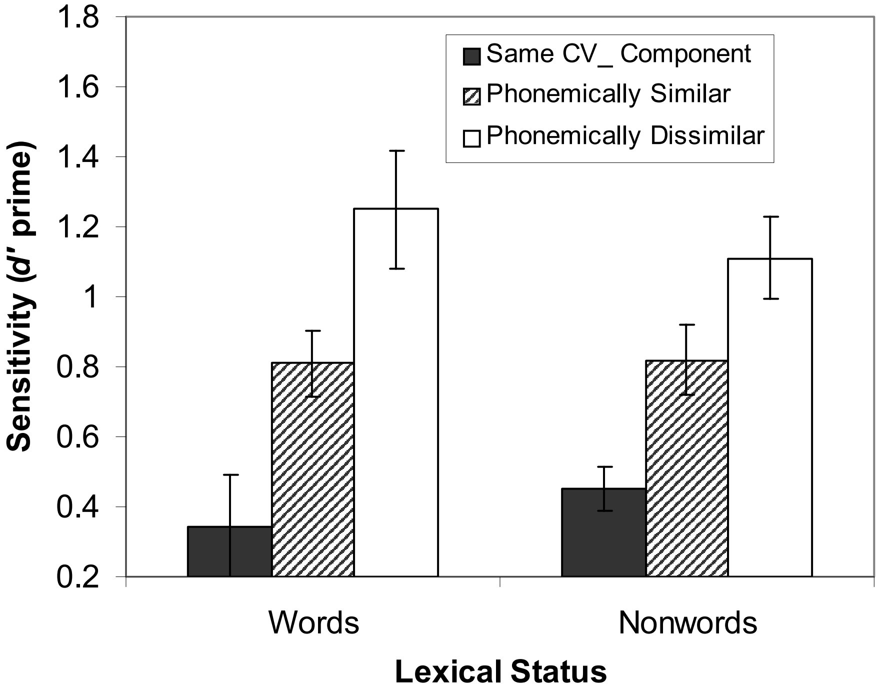

Figure 6.2 Means (± SE) for phonemically similar, phonemically dissimilar and CV_ lists as a function of lexicality for Experiment 2, using a measure of sensitivity (*d'* prime).

6.4.3. Discussion

Regardless of whether the stimuli were words or nonwords, phonemic similarity had a detrimental effect on order memory, in that order memory was worse when the stimulus lists shared the CV_ component as compared to phonemically similar lists, which was worse than when the lists were phonemically dissimilar. These findings are consistent with STM models that assume that as phonemic similarity increases order memory decreases (non-linguistic STM models; Nairne, 1988, 1990a), and with the idea that order memory is impaired to a greater extent when list items share a common vowel (psycholinguistic models of STM; Gupta and MacWhinney, 1997; Hartley & Houghton, 1996).

Further, the pattern of results observed in the current study is identical to the correct-in-position results we (Study 1; Experiment 2) found with words using an immediate serial recall task. As such, the current findings strengthen the argument that when the experimental stimuli are nonwords, speech production processes influence the results obtained from tasks that require participants to verbally recall list items (e.g., the serial recall task).

Finally, the current study found no lexicality effect for order recognition judgements. In other words, a participant's ability to recognize that a list was in the 'same' or a 'different' order was similar, regardless of whether the stimulus lists were words or nonwords. This finding is consistent with the idea that the lexicality effect should be evident whenever the task requires the retention of item information and attenuated or absent when the task does not (Gathercole et al., 2001).

6.5. Experiment 3

Experiment 3 was again designed to further examine the effect of phonemic similarity on order memory once the demands that overt speech production have on STM performance are removed. Experiment 3 consisted of lists of words and nonwords that shared the C_C component.

More importantly, the current experiment also aimed to critically examine the predictions that non-linguistic and psycholinguistic models of STM generate for the effect of phonemic similarity on order memory. For instance, non-linguistic STM models (Nairne, 1988, 1990a) predict that as similarity increases order memory should decrease. Hence, according to these types of models, order memory should be poorer for C_C as compared to the phonemically similar lists. In contrast, psycholinguistic models of STM (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) suggest that any form of phonemic similarity between list items should have a detrimental effect on order memory (i.e., C_C lists) but the greatest impairment should be observed when the list items share a common vowel (i.e., the phonemically similar lists in the current experiment). It remains to be seen whether the similarity of sharing both consonants will outweigh sharing the vowel, however, in a serial recall task we found that order memory when measured using the correct-in-position criterion was better for C_C as compared to phonemically similar lists of words (i.e., Study 1; Experiment 3).

6.5.1. Method

6.5.1.1. Participants

Seventy-two undergraduate psychology students from the University of Wollongong participant pool (16 males and 56 females), with an age range from 16 to 49 years (M = 21), participated in compliance with a course requirement. The same inclusion criteria were placed on the selection of participants for the current experiment as in Experiment 1.

6.5.1.2. Stimuli

Two stimulus sets composed of either words or nonwords were constructed.

Word pool. The word pool comprised 180 words with a CVC structure (refer Appendix A – Table A3). The stimuli were used to create 30 same-consonant (C_C), 30 phonemically similar and 30 phonemically dissimilar six-word lists. The same constraints were placed on the construction of the stimulus lists as for Experiment 1, except for two minor modifications. The first modification was that for the same consonant condition, all of the stimuli in a particular list shared the C_C component (e.g., *Bought, Bet, Boot, But, Bat,* and *Bait*). Also, for the phonemically similar condition, no items in a list shared both consonants (e.g., *Bought, Bored, Lawn, Wrought, Fort,* and *Fawn*). Hence, all of the items in a phonemically similar list shared the same vowel.

Nonword pool. The nonword pool comprised 180 nonwords with a CVC structure (refer Appendix E – Table E3). The same constraints were place on list construction for the nonwords as outlined above.

6.5.1.3. Procedure

The same testing procedure was used in the current experiment as in Experiment 1.

6.5.2. Results

The same analyses were performed on the data obtained in the current experiment as in Experiment 1. The mean proportions of correct responses were collated for words and nonwords when the stimuli were phonemically dissimilar, similar or shared the C_C component (refer Table 6.3).

Table 6.3 *Mean Proportions (with Standard Deviations) of Lists Correctly Recognized in Experiment 3 as a Function of Lexicality and List Type.*

	C_C lists		*Phonemically Similar*		*Phonemically Dissimilar*	
Lexicality	*M*	*SD*	*M*	*SD*	*M*	*SD*
Words	0.66	0.11	0.61	0.10	0.73	0.11
Nonwords	0.62	0.11	0.59	0.08	0.68	0.10
Mean	0.64		0.60		0.70	

Figure 6.3 summarises the mean *d'* values for words and nonwords when the stimuli were phonemically dissimilar, similar or shared the C_C component. The *d'* analysis revealed a main effect of phonological similarity, $F(2,140) = 28.434$, $MSE = .321$, $p < .01$. The analysis also revealed a main effect of lexicality, $F(1,70) = 5.005$, $MSE = .726$, $p < .05$, in that order memory was better for words ($M = .992$) as compared to nonwords ($M = .729$). Phonological similarity was not found to interact with

lexicality, $F(2,140) = 1.844$, *ns*. Post-hoc analyses revealed that order memory was better when the stimuli were phonemically dissimilar as compared to lists that shared the C_C component, $t(29) = 5.176$, $p < .0167$, which was better than when the stimuli were phonemically similar, $t(29) = 2.458$, $p < .0167$(although no differences between the C_C and phonemically similar conditions were found when the analysis was performed using the mean proportions of correct responses).

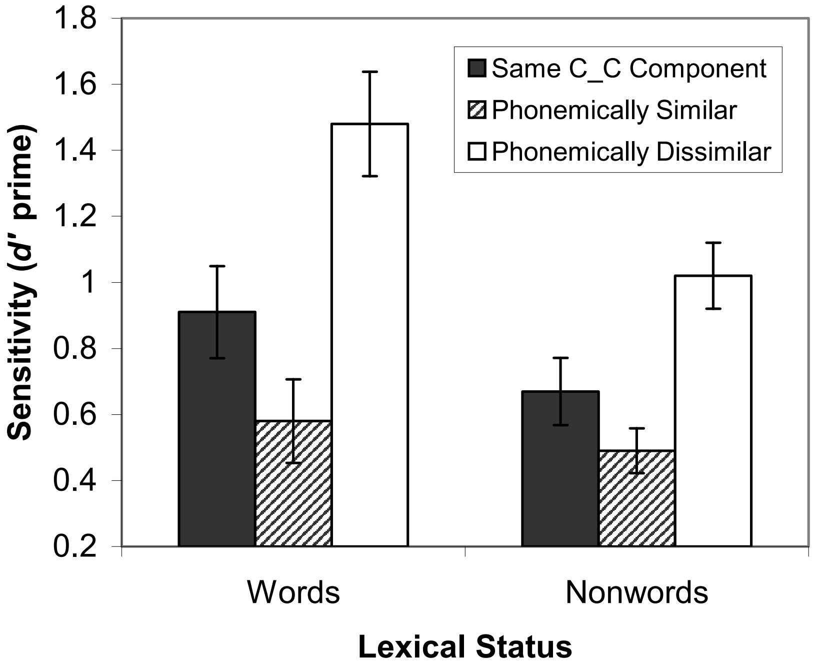

Figure 6.3 Means (± SE) for phonemically similar, phonemically dissimilar and C_C lists as a function of lexicality for Experiment 3, using a measure of sensitivity (*d'* prime).

6.5.3. Discussion

Regardless of whether the stimuli were words or nonwords, phonemic similarity had a detrimental effect on order memory, in that order memory was better for phonemically dissimilar lists than stimulus lists that shared the C_C component, which was better than when the stimulus lists were phonemically similar. These findings are identical to the results we (Study 1; Experiment 3) found when the experimental stimuli were words and order memory was measured using the correct-in-position criterion. However, we (Study 2; Experiment 3) found a different pattern of results when the experimental stimuli were nonwords, in that correct-in-position performance was better for C_C as compared to when the stimuli were phonemically dissimilar. As there was no interaction between lexicality and similarity, the results of the current study lend strong support, not only to the idea that words and nonwords are processed in a similar way (Fallon et al., in press), but also to the idea that the results obtained from studies that require participants to overtly recall lists of nonwords (i.e., Study 2) are influenced by the processes involved with speech production.

More importantly, the findings from the current experiment are inconsistent with the predictions generated from non-linguistic models of STM (Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a). According to STM models that are based on the distinctiveness assumption, as similarity increases order memory should decrease. However, in the current experiment, lists of C_C items shared a larger number of overlapping phonemes than did the phonemically similar lists.

The finding of an order memory advantage for C_C as compared to the phonemically similar lists is, however, consistent with the predictions generated by psycholinguistic models of STM (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996). For instance, according these types of models, any form of phonemic similarity should have a detrimental effect on order memory (i.e., phonemically dissimilar as compared to C_C lists) but the greatest impairment will be seen when the overlapping phoneme is the vowel (i.e., in this case, the phonemically similar condition).

Finally, contrary to the findings observed in Experiments 1 and 2, the current study found a main effect of lexicality. In other words, order memory was better when the experimental stimuli were words as compared to nonwords. We are uncertain why this occurred. One suggestion that may explain these findings stems from the idea that the lexicality effect should be observed whenever a task requires that participants remember item information. For instance, when participants are presented with lists of _VC (Experiment 1) or CV_ (Experiment 2) items, an effective strategy that participants may use to retain the order in which the list items are presented would be to 'remember the initial or final consonant'. For example, all a participant needs to encode is the final consonant (i.e., *p, n, f,* or *b, m, g*) to remember the order in which the list items *wipe, wine, wife* or *wieb, wiem, wieg,* were presented. The same applies to lists that rhyme, except that in this case an effective strategy would be to remember the initial consonant. However, when the lists are composed of C_C stimuli, 'remember the vowel' may not be an effective strategy (e.g., *house, hiss, horse or hars, hays, hies*). As such, these types of lists (i.e., shared C_C component) may force participants to encode all of the item information as compared to the use of a partial information strategy that is possible when lists share either the CV_ or _VC components. Hence, the lexicality effect observed in the current study may reflect a larger role of item information in this experiment. As such, future research may need to design a purer serial recognition task that forces participants to encode all of the item information presented in a list.

6.6. General Discussion

In summary, regardless of how phonemic similarity was operationally defined, the results were consistent with a recent study conducted by Gathercole et al. (2001), in that order memory was better for phonemically dissimilar lists than stimulus lists that shared any form of phonemic overlap. Hence, the findings from the current study are consistent with the suggestion that the effect of phonemic similarity on order memory persists once the demands that overt speech production has on STM performance are removed.

Although some researchers suggest that the same processes are involved in word and nonword recall (e.g., Fallon et al., in press), others have argued that different STM processes are involved (e.g., Brown & Hulme, 1995). We have recently found different similarity effects on correct-in-position recall that were dependent on whether the experimental stimuli were words (Study 1) or nonwords (Study 2). Therefore, the main impetus for conducting the current research was to examine whether the performance differences observed across the above-mentioned studies were due to differences in the STM processes that are operating when the stimuli are words as compared to nonwords. The current research findings go a long way to clarify this issue. For instance, varying the stimulus type (i.e., words or nonwords) while keeping the amount of phonemic overlap constant across experiments produced an identical pattern of results. Hence, the results observed in the current study lend strong support to the idea that similar STM processes are involved in both word and nonword recall (Fallon et al., in press).

Further, regardless of whether the stimuli were words or nonwords, the findings from the current research perfectly replicated those obtained in the word study (Study 1). Hence, these findings suggest that speech production processes influence the results obtained from studies that require participants to overtly articulate lists of nonwords (Gathercole et al., 2001). Therefore, the current findings suggest that the differential results observed across studies that have assessed word and nonword recall with tasks that require the verbal recall of list items are a product of speech production processes that constrain STM performance when the experimental stimuli are nonwords, rather than to differences in the STM processes involved when recalling these types of stimuli.

6.6.1. Implications for short-term memory models

The findings from the current study also have implications for models of STM. A further aim of the current study was to examine whether the effect of phonemic similarity on order memory is due to the distinctiveness of list items in relation to the other list items presented (Nairne, 1988, 1990a), or to linguistic processes that are operating at the sub-syllabic level (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996). This is important in that the effect that sub-syllabic linguistic mechanisms have on STM performance has previously only been demonstrated with the serial recall task which requires the overt articulation of presented list items (Studies 1 & 2).

In line with previous research (Studies 1 & 2), the findings from the current study are clearly consistent with the predictions generated by psycholinguistic models of STM (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996). Regardless of whether the stimuli were words or nonwords, phonological similarity had a detrimental effect on order memory, however, when the overlapping phoneme was the vowel (i.e., _VC lists, Experiment 1[15]; CV_ lists, Experiment 2; phonemically similar lists, Experiment 3), a further order memory impairment was observed. Therefore, the current findings argue against STM models that rely on the distinctiveness or 'phonemic overlap' assumption

[15] Although non-significant, the findings were in the direction predicted with poorer order memory for rhyming as compared to phonemically similar lists.

in which to explain the effect of phonological similarity on order memory (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a).

To reiterate, the current finding of a sub-syllabic influence on order memory, with an STM task that is independent of overt speech production, lends further support to the claim that the influence of sub-syllabic mechanisms on STM performance is a genuine memory effect and not simply due to speech production processes or factors such as the ease of the articulatory transitions between lists items (see Murray & Jones, 2002). As such, future STM models will need to incorporate the mechanisms necessary to deal with linguistic processes that are operating at the sub-syllabic level to constrain STM performance.

7. A GENERAL SUMMARY OF THE CURRENT RESEARCH FINDINGS

7.1. A summary of the current research findings

Across three studies, phonemic similarity was manipulated by operationally defining similarity in different ways. Lists of items rhymed, or shared either the CV_ or C_C components. This varied the position of the overlapping phonemes between the phonemically similar conditions, while keeping the amount of phonemic overlap constant. For all three studies, STM performance was compared with lists that were either phonemically similar (i.e., each stimulus in each list consisted of at least two phonemes in common with at least one other stimulus in the same list) or phonemically dissimilar (i.e., each stimulus in each list did not share any phonemes in common with any other stimulus in the same list).

Regardless of whether the stimulus lists were words or nonwords, STM performance was assessed using either a serial recall or serial recognition task. When the task was serial recall, item recall (i.e., scored as correct if a participant recalled a presented list item, regardless of position), correct-in-position (i.e., scored as correct if a participant recalled the correct item in the correct position) and order accuracy (i.e., the score obtained using the correct-in-position criterion divided by the score obtained using the item recall criterion) measures of performance were obtained. To take account of potential shifts in bias, or levels of attention when the task was serial recognition, hits and false alarm rates were used to calculate *d'* prime values across each of the conditions. The following sections summarise and compare the results obtained across the three studies. To aid in this discussion, a summary of the main findings are presented in Table 7.1.

Table 7.1 Summary of the Pattern of Research Findings across the Three Studies when the Stimulus Lists were Either Words or Nonwords, and the Task was Either Serial Recall or Serial Recognition.

Research Findings	**Words**	**Nonwords**
Rhyming Experiments		
Serial Recall		
Item Recall	Rhyming > Similar = Dissimilar	Rhyming > Similar > Dissimilar
Correct-in-position Recall	Dissimilar > Similar = Rhyming	Rhyming = Similar > Dissimilar
Order Accuracy	Dissimilar > Similar > Rhyming	Dissimilar > Similar > Rhyming
Serial Recognition	Dissimilar > Similar = Rhyming	Dissimilar > Similar = Rhyming
CV_ Experiments		
Serial Recall		
Item Recall	Similar = CV_ = Dissimilar	Similar > CV_ > Dissimilar
Correct-in-position Recall	Dissimilar > Similar > CV_	Similar > CV_ = Dissimilar
Order Accuracy	Dissimilar > Similar > CV_	Dissimilar > Similar > CV_
Serial Recognition	Dissimilar > Similar > CV_	Dissimilar > Similar > CV_
C_C Experiments		
Serial Recall		
Item Recall	C_C = Dissimilar > Similar	C_C > Similar > Dissimilar
Correct-in-position Recall	Dissimilar > C_C > Similar	C_C > Similar = Dissimilar
Order Accuracy	Dissimilar > C_C = Similar	Dissimilar > C_C > Similar
Serial Recognition	Dissimilar > C_C > Similar	Dissimilar > C_C > Similar

7.2. Operationally defining phonological similarity

For each of the experiments in any one study, phonological similarity was operationally defined in a number of different ways. For the rhyming experiments, similarity was defined in terms of lists that shared a rhyme ending, and compared to performance when the stimulus lists were either phonemically similar (i.e., no items in a list shared a rhyme ending and each stimulus in each list consisted of at least two phonemes in common with at least one other stimulus in the same list) or phonemically dissimilar (i.e., none of the items in a list shared any common phonemes). For the CV_ experiments, similarity was defined in terms of lists that shared the CV_ component, and compared to performance when the stimuli were either phonemically dissimilar or phonemically similar (i.e., the same constraints were placed on construction of list items for the similar conditions in these experiments except that, none of the items in these lists shared the CV_ component). Finally, similarity for the C_C experiments was defined in terms of lists that shared the C_C component. Performance was again compared to either a baseline performance measure (i.e., dissimilar lists), or to performance when the stimuli were phonemically similar (i.e., none of the items in these lists shared the C_C component).

7.3. Aims of research using the immediate serial recall task with words

There were two main aims for conducting study one. Current research suggests that phonological similarity differentially influences order memory and the recall of item information, in that similarity impairs order memory but has been found to facilitate the recall of item information (Fallon et al., 1999). Therefore, a main aim of this study was to examine the effect that phonemic overlap has on both item and order memory when phonemic similarity was operationally defined in different ways. A further aim of this study was to critically evaluate the explanations generated by STM models for the effect that similarity has on the recall of item information, and memory for an item's position within a list.

7.3.1. Conclusions drawn from the serial recall studies with words

There are two issues that research into the PSE needs to address: the effect that similarity has on order memory, and its effect on the recall of item information. At the order recall level, the findings from the current experiments (i.e., using either correct-in-position or order accuracy measures) suggest that any form of phonemic similarity impairs order memory (see Table 7.1). However, when performance was measured using the correct-in-position criterion, the current study found that order memory was worse for the phonemically similar as compared to C_C lists (Experiment 3). This is important in that the C_C lists shared a greater number of overlapping phonemes than did the phonemically similar lists. As such, this finding is inconsistent with the predictions generated from STM models that are based on the ratio rule (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a). In contrast, the Hartley and Houghton (1996) model suggests that sharing a number of overlapping phonemes should impair order memory. However, when the overlapping phoneme is the vowel, as was the case for the rhyming (Experiment 1), CV_ (Experiment 2) and phonemically similar lists (Experiment 3), order memory should be further impaired. Hence, the findings observed in the current study are consistent with the idea that, the influence that similarity has on order memory is dependent on the phonemic make up of the presented list items.

Currently, two different accounts have emerged in the research literature that provide an explanation for the facilitative effect that sharing a rhyme ending has on the recall of item information. For instance, according to current versions of the feature model, retrieval cues such as a shared rhyme ending can be used to limit the size of the secondary memory search set (Nairne, 2002; Neath, 1999). However, if limiting the size of the secondary memory search set can be used to facilitate the recall of item information, then stimulus lists that share features that make them easily discriminable along the list dimension (i.e., CV_ or C_C lists) should also facilitate item recall. This prediction is clearly at odds with the current research findings of no differences in the recall of item information between the phonemically dissimilar as compared to either C_C (Experiment 3) or CV_ lists (Experiment 2). Alternatively, Hartley and Houghton (1996) suggest that sub-syllabic structures aid in the recall of item information. For instance, if a syllable is thought of in terms of an item with a distinct onset-rhyme structure, then list items that share this structure can serve as a prime to aid the recall of item information (Hartley & Houghton, 1996). When list items do not rhyme, syllable representations do not receive the same level of support. Hence, in comparison to rhyming lists, the recall of item information should not benefit to the same extent when stimulus lists share either the CV_ or C_C components. Therefore, the findings observed in the current study are more consistent with the explanations that Hartley and Houghton's (1996) linguistically constrained model of STM generate for the effect that phonemic similarity has on the recall of item information, and memory for an item's position in a list (see Table 7.1).

In summary, the findings from the current study rule out the possibility that the item recall advantage observed for rhyming lists of items is due to phonemic overlap. Further, the findings suggest that it is no longer sufficient to use a distinctiveness argument to explain the effect that similarity has on order memory. Rather, the results suggest that performance on STM tasks is influenced by linguistic mechanisms that operate at the sub-syllabic level. Therefore, to adequately explain the current research findings, future STM models need to incorporate the mechanisms necessary to deal with the psycholinguistic rules that constrain STM performance at the sub-syllabic level.

7.4. Aims of research using the immediate serial recall task with nonwords

There were three main aims for examining the effect that similarity has on the recall of nonwords using the ISR task. Firstly, if words and nonwords are processed using the same STM mechanisms, then the findings from the current study should mirror those observed in study one. The second aim of the current study was to assess the utility of STM models by the explanations they generate for the effect that phonemic similarity has on the recall of item information, and memory for an item's position in a list when the experimental stimuli are nonwords. Finally, to our knowledge the current study was the first direct test of the idea that, when the experimental stimuli are phonemically similar nonwords, linguistic constraints, such as syllable structure and sonority, operate at the sub-syllabic level to influence STM performance.

7.4.1. Conclusions drawn from the serial recall studies with nonwords

The findings from the current study are inconsistent with the predictions based on newer versions of the feature model (i.e., Nairne, 2002; Neath, 1999). For instance, this model predicts an item recall advantage for lists that share features that can be used to limit the size of the secondary memory search set (i.e., rhyming, CV_ or C_C lists). However, the current study did not find an item recall advantage for CV_ as compared to phonemically similar lists of nonwords (see Table 7.1). Further, the findings from the current study are also inconsistent with Hartley and Houghton's (1996) linguistically constrained model of STM. According to this model, the recall of item information should not benefit to the same extent when the stimulus lists share either the CV_ or C_C components as compared to when the stimulus lists rhyme. However, when compared to the phonemically dissimilar conditions, the current study found an item recall advantage not only for rhyming lists of nonwords (Experiment 1), but also for nonword lists that shared the C_C component (Experiment 3). Finally, across all three experiments, the recall of item information was worse when the experimental lists were composed of phonemically dissimilar nonwords, regardless of how similarity was operationally defined (refer Table 7.1).

However, a note of caution should be exercised when drawing conclusions from the results observed when performance was scored using the item recall criterion. Firstly, item recall performance across all three of the phonemically dissimilar conditions was extremely low. For instance, participants were recalling on average only 1 to 1.5 items out of five. Thus, participants were only able to remember the first (or last) item in each list. This finding makes sense in that when the experimental stimuli are phonemically dissimilar nonwords there is nothing to facilitate the recall of item information (e.g., a shared rhyme ending or repetition priming when the experimental stimuli are phonemically similar). Secondly, when there are high levels of item recall (i.e., as observed in Study 1), item scores produce reliable, meaningful and lawful results.

When performance was measured using the correct-in-position criterion, two different patterns of results emerged (refer Table 7.1). For instance, in comparison to when the experimental lists were composed of phonemically dissimilar nonwords, an order memory advantage was observed when the stimuli either rhymed (Experiment 1) or shared the C_C component (Experiment 3). Whereas, when the nonwords shared the CV_ component (Experiment 2), in comparison to the phonemically dissimilar lists, no difference in order memory was observed (see Table 7.1). However, the results observed when order memory is scored using the correct-in-position criterion are not independent from the recall of item information. This is further demonstrated by the contradictory findings observed in previous studies that have used the correct-in-position criterion to examine the effect that phonemic similarity has on order memory (e.g., Besner & Davelaar, 1982; Drewnowski, 1980; Fallon et al., in press; Lian et al., 2001). As such, due to the influence that differences in item recall ability have on the correct-in-position measure of performance, a clear picture of the influence that similarity has on order memory cannot be gained by using this measure.

However, once the influence that individual differences in item recall ability has on order memory is controlled by using the order accuracy measure of performance, the results obtained from both the word (Study 1) and nonword (Study 2) experiments correspond almost perfectly (see Table 7.1). Further, comparison across the nonword experiments revealed that order memory was better for C_C (Experiment 3) as compared to either rhyming (Experiment 1) or CV_ (Experiment 2) lists. Therefore, as with the results observed for the word study (Study 1), the current findings suggest that it is not the degree of phonemic overlap, or the distinctiveness of a particular item in relation to all of the other list items that influences order memory, *per se*, but the consistency and phonemic make up of the list items. Further, these findings strengthen the argument mounted by Fallon et al. (in press) that words and nonwords are processed in the same way.

In summary, the current findings suggest that the same mechanisms are involved in both word and nonword recall. Secondly, as with words (Study 1), explanations for the effect that similarity has on order memory that rely on the 'phonemic overlap' or distinctiveness argument, are no longer sufficient. Finally, the findings from the current study (as with Study 1) suggest that linguistic constraints, such as syllable structure and sonority, which are assumed to operate at the sub-syllabic level, influence both the recall of item information and memory for an item's position in a list. Therefore, to provide an explanation for the current research findings, future STM models will need to include linguistic constraints on STM performance at the sub-syllabic as compared to lexical level.

7.5. Aims of research into the phonological similarity effect using the serial recognition with both words and nonwords

There were three main aims for conducting the current study (i.e., Study 3). The first aim was to assess whether the effect that operationally defining phonemic similarity in different ways has on order memory, observed in the previous experiments (i.e., Studies 1 and 2), remains, once the demands that overt speech production have on STM performance are removed. The second aim of study three was to examine whether the effect that phonemic similarity has on order memory is similar, regardless of whether the experimental stimuli are words or nonwords. The final aim of the current study was to critically examine the predictions that non-linguistic (Nairne, 1988, 1990a) as compared to psycholinguistic (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) models of STM generate for the effect that phonemic similarity has on order memory.

7.5.1. Measuring short-term memory performance when the task is serial recognition

Traditionally, there are two ways in which performance can be measured when the task is serial recognition. One method is to sum the number of correct responses made by a particular individual (i.e., the number of times a participant says ‘same’ when the items were in the same order, and the number of times a participant says ‘different’ when the items were in a different order). However, this method of measuring performance does not take into account a particular individual’s bias toward responding in either one way or the other.

There are two ways in which recognition scores can be compared. One of the simplest techniques for dealing with the problem of false alarms is to use a *guessing correction* (Baddeley, 1997). The guessing correction derives from the assumption that, if participants remember that lists of items were in the same order as previously presented lists, then they will correctly categorise the lists as ‘same’. However, for the remaining lists, if participants do not remember the order, they will guess (Baddeley, 1997). If the same numbers of lists were in the 'same' as compared to 'different' order, a participant's odds of correctly classifying a list through guessing would be 50%. As such, every time a participant says 'different' when the lists were in the same order, given the odds of correctly classifying a list by guessing, there should also be one list that a participant has correctly classified as 'same', that can be attributed to guessing. Hence, a participants sensitivity to a particular recognition task can be calculated "…by subtracting the number of false alarms from the number of correct detections" (Baddeley, 1997; p.197).

An alternative way of dealing with false alarms derives from *signal detection theory* (McNicol, 1972). From a practical point of view, signal detection theory provides a number of useful performance measures when participants are required to make a decision. According to signal detection theory, there are two aspects of performance that are traditionally confounded: *sensitivity* and response *bias*. The measure of how well an individual is able to make a correct judgment and avoid an incorrect judgment has been termed *sensitivity.* In contrast, *bias* refers to the idea that participants may favour one response over the other (McNicol, 1972). In comparison to the *guessing correction* which controls for task sensitivity, signal detection theory is able to tease apart the effect that response bias has on STM performance, from a participants sensitivity to a particular task (McNicol, 1972). Hence, in line with previous research that has used the recognition paradigm to study STM (Banks, 1970; Lockhart & Murdock, 1970), the current thesis has used d-prime (*d'*) as a measure of sensitivity.

7.5.2. Conclusions drawn from serial recognition studies with words and nonwords

Although the results obtained across the experiments varied as a consequence of how phonemic similarity was operationally defined, the influence that similarity had on order memory was the same, regardless of whether the stimuli were words or nonwords (refer Table 7.1). Further, the findings from the current study are once again inconsistent with the predictions generated by non-linguistic STM models that are based on the ratio rule (e.g., Brown et al., 2000; Burgess & Hitch, 1992; 1999; Nairne, 1990a). For instance, although order memory was better for phonemically similar as compared to CV_ lists (Experiment 2), no differences were found between the phonemically similar and rhyming lists of stimuli (Experiment 1). Further, in Experiment 3, order memory was worse for the phonemically similar as compared to stimulus lists that shared the C_C component. These findings replicate the results obtained across all three experiments when the experimental stimuli were words, and order memory was measured using the correct-in-position criterion (i.e., Study 1).

Given the suggestion that the serial recognition task is not as sensitive to experimental manipulations as a task such as serial recall (Gathercole et al., 2001), one could argue that the finding of no order memory difference between the phonemically similar and rhyming[16] lists of stimuli (Experiment 1), is due to a lack of sensitivity. Furthermore, one could argue that given enough power, this finding may in fact be consistent with the predictions generated by non-linguistic (i.e., as similarity increases order memory should decrease) and psycholinguistic (i.e., any form of similarity has a detrimental effect on order memory, but when the overlapping phoneme is the vowel, a further impairment should be observed) models of STM. Therefore, measures of effect size were obtained to examine whether the null finding observed between phonemically similar and rhyming lists of stimuli when the task was serial recognition (Experiment 1) was due to a lack of power. When the task was serial recall and performance was measured using the correct-in-position criterion*,* the epsilon-squared statistic for the main effect of similarity was larger (E_R^2 = .4539) than the effect size obtained when the task was serial recognition (E_R^2 = .1664). This finding lends support to the idea that the measure of order memory obtained when the task is serial recognition may not be as sensitive a measure of performance as is obtained when the task is serial recall. Further, the findings were in the direction predicted by non-linguistic (Nairne, 1988, 1990a) and psycholinguistic (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) models of

[16] Although, no difference in order memory performance was observed between the phonemically similar and rhyming lists of words when measured using the correct-in-position criterion (i.e., Study 1; Experiment 1), when scored using a purer measure of order memory (i.e., order accuracy), order memory was worse for rhyming as compared to the phonemically similar condition.

STM, in that order memory was worse for rhyming as compared to when the stimuli were phonemically similar.

Hence, as with the previous studies (i.e., Studies 1 & 2), the findings from the current experiments are inconsistent with the idea that the influence that phonemic similarity has on order memory is due to the *distinctiveness* of list items in relation to other list items (e.g., Brown et al., 2000; Burgess & Hitch, 1992; 1999; Nairne, 1990a). Rather, the findings lend support to the suggestion that order memory is influenced by linguistic constraints, such as sonority, that operate at the sub-syllabic as compared to lexical level. As such, the findings observed in the current study are more consistent with the predictions generated from psycholinguistic models of STM (Gupta & MacWhinney, 1997; Hartley & Houghton, 1996).

Further, the findings from the serial recognition experiments are inconsistent with the idea that different processes are involved in word and nonword recall (Lian et al., 2001; Saint-Aubin & Poirier, 2000). Rather, these findings lend support to the suggestion that words and nonwords are processed in a similar way (Fallon et al., in press). Also, the fact that the results observed in the current study match those obtained when order memory was measured using the correct-in-position criterion (i.e., Study 1) suggest that when the task is serial recall, overt speech production processes influence STM performance to a greater extent when the stimuli are nonwords as compared to words (Gathercole et al., 2001). Finally, these findings suggest that once the demands that overt speech production have on STM performance are removed, the influence that similarity has on order memory, at the sub-syllabic level, remains.

7.6. Putting the research findings obtained in the current thesis into perspective

To reiterate, a number of clear conclusions can be drawn from the research findings observed in the current thesis. Firstly, the findings from the current thesis suggest that the same mechanisms are involved in both word and nonword recall. Thus, strengthening the claim that the influence that similarity has on STM performance is at the sub-syllabic as compared to lexical level. Secondly, the results observed when a serial recognition as compared to a serial recall task is used to measure performance suggest that the effect that phonemic similarity has on order memory persists once the demands that overt speech production have on STM performance are removed. Also, the findings observed in the current thesis suggest that overt speech production processes influence the results obtained when using the serial recall task, especially when the experimental stimuli are nonwords.

Finally, the current findings have larger implications in that the majority of STM models use a 'phonemic overlap' or distinctiveness assumption to explain the effect that similarity has on STM performance. However, the findings from the current thesis suggest that STM models that are based on the ratio rule (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a), cannot adequately deal with the current research findings. Rather, the results suggest that, when the experimental stimuli are phonemically similar, both the recall of item information and memory for an item's position in a list are influenced by linguistic constraints that are operating at the sub-syllabic level. Hence, if STM models are to provide a successful account of the current research findings, it is imperative that they incorporate the linguistic constraints, such as syllable structure and sonority that influence STM performance at the sub-syllabic as compared to lexical level.

8. IMPLICATIONS AND FUTURE RESEARCH DIRECTIONS

"The phonological similarity effect has achieved the status of a 'benchmark' finding in the immediate memory literature, and most theories of short-term memory include mechanisms that are specifically designed to account for the phenomenon" (Nairne & Kelley, 1999; p.45). As Gathercole (1997; see also Page & Norris, 1998) suggests, one value of STM models lies in their ability to account for the effect that phonemic similarity has on STM performance. However, the results from the current research all lead to the same suggestion: extant STM models *cannot* adequately deal with the current research findings. For instance, the majority of STM models are based on the idea that it is the distinctiveness of list items in relation to other list items that is important for order memory (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a; Tehan & Fallon, 1999). However, in combination, the results from the current thesis strongly suggest that regardless of whether the stimuli are words or nonwords, or whether or not the task requires participants to verbally reproduce the presented items, linguistic mechanisms operate at the sub-syllabic level to influence STM performance.

Given the importance of the PSE for STM modelling, the following discussion will concentrate on the implications that the current research findings have for extant STM models. For instance, 'Are slight modifications to existing STM models able to account for the effect that similarity has on STM performance?' or 'Are more plausible explanations derived from models that are based on research within another domain?' Although modifications to any number of STM models could be proposed in an attempt to account for the current research findings, Nairne's (1988, 1990a, 2002) feature model of immediate memory was selected as the exemplar for other STM model that explain STM performance in terms of general principles. Further, given that the findings from the current research suggest that mechanisms are operating at the sub-syllabic as compared to lexical level to influence STM performance, Glasspool's (1995) model of STM that accounts for lower order effects (i.e., sequential ordering of phonemes) will be examined to assess whether slight modifications to existing STM models have the mechanisms necessary to deal with the current research findings.

Finally, an easy argument to mount is that any scientific endeavour is only as good as the future research it generates. Therefore, future research directions that have been generated from the current research will be discussed in detail.

8.1. Nairne's feature model of immediate memory revisited

As suggested previously, most STM models attempt to provide an explanation for the effect that phonological similarity has on STM performance at the lexical level (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999). However, one model that explains STM performance in terms of processes and principles that are applicable to almost any type of information (e.g., objects, spatial location, pictures or verbal material) is Nairne's (1988, 1990a, 2002) feature model of immediate memory.

8.1.1. Factors that affect the utility of trace features for recall

According to the feature model, there are two factors that influence the utility of trace features for recall: distinctiveness and salience. Distinctiveness refers to the phonological characteristics of a particular item and an item's phonemic similarity to its closest neighbour in a presented list. In other words, the likelihood that a particular list item will be recalled is a function of the relative similarity of the probe item to all of the presented list items (Gillund & Shiffrin, 1984; Hintzman, 1986; Luce, 1959; Nosofsky, 1986). Hence, STM models that are based on the distinctiveness assumption predict that as phonemic similarity increases order memory should decrease. However, the feature model (1988, 1990a, 2002) assumes that the likelihood of recalling a particular list item is also influenced by the salience of the trace features. According to Nairne (1988), cues can be used to increase the salience of list items to the extent that participants perceive the items as belonging to the same category.

8.1.2. Current research findings in relation to Nairne's (1988, 1990a, 2002) feature model of immediate memory

If STM models are to successfully account for the effect that phonemic similarity has on STM performance, they need to explain not only the effect that phonemic similarity has on order memory, but also the effect that it has on the recall of item information.

In terms of the effect that phonemic similarity has on order memory, the feature model (Nairne, 1988, 1990a) is based on the distinctiveness assumption and therefore, cannot provide a plausible explanation for the effect that phonemic similarity has on order memory. However, this model also assumes that the salience of list items is important for recall. According to Hartley and Houghton (1996) the speech trace for vowels is both longer in duration and more acoustically intense in comparison to the speech trace for consonants. Hence, the parameters of the feature model (Nairne, 1988, 1990a) could be modified to reflect the greater salience of the speech trace for vowels as compared to consonants. This slight modification, where the vowel is given a heavier weighting than consonants, would allow the feature model to account for the effect that similarity has on order memory for the same reason that Hartley and Houghton's (1996) linguistically constrained model of STM for nonwords can.

In terms of the effect that phonemic similarity has on the recall of item information, newer versions of the feature model (Nairne, 2002) assume that an item recall advantage should be observed whenever the 'secondary memory search set' can be limited to a smaller number of items. In other words, cues can be used to increase the salience of list items to the extent that participants perceive the items as belonging to the same category. Therefore, it is possible to argue that the rhyme unit may be a more salient cue than other list cues, and as such provide an explanation for the item recall advantage observed for rhyming lists of items. However, to make a cohesive argument, this model would need to specify why this may be the case in comparison to lists that share either the CV_ or C_C component.

8.2. An extension of the Burgess and Hitch (1992) model (Glasspool, 1995)

Another STM model that does not rely purely on mechanisms that are assumed to operate at the lexical level was developed by Glasspool (1995). Glasspool (1995) extended the Burgess and Hitch (1992) model to include a mechanism that is able to serially order items at the phoneme level. Hence, the Glasspool (1995) model, with items being represented at the phoneme level, is able to model the effect that phonemic similarity has on STM performance when the experimental stimuli are nonwords.

8.2.1. The serial ordering of phonemes within words (Glasspool, 1995)

Glasspool (1995; p.14) modified the Burgess and Hitch (1992) model so that the phonemes that an item is comprised of could be temporally sequenced. As Glasspool (1995) suggests, the "…phonemes constituting words and nonwords should be sequenced by the same mechanism as that which sequences words within lists". Hence, this model incorporates a CQ mechanism, as described by the Burgess and Hitch (1992) model (refer to section 2.3.3.1 for a description of the basic CQ architecture), that is responsible for temporally ordering the phonemes within presented list items.

According to Glasspool (1995), there are two memory systems: a memory system for words and a memory system for phonemes (see Figure 8.1). Both systems are assumed to consist of a competitive filter and a time-varying context signal (refer section 2.3.3.3). When an item is presented to the memory system, this system attempts to remember the item, regardless of whether it is a word or nonword (Glasspool, 1995). In other words, phoneme nodes are activated irrespective of whether the presented item is a word or nonword. In this model, the lexicality effect derives from the idea that, "…additional support from lexical information in long-term memory over and above a basic phonological capability" aids recall attempts when the stimuli are words (Glasspool, 1995; p. 15). In contrast, when the stimuli are nonwords, although access to the word system is still attempted during list presentation, this word system is unable to support the phoneme system during recall.

Please see print copy for Figure 8.1

Figure 8.1. An outline of Glasspool's (1995) STM model.

Basic to all STM models that have incorporated a context-timing signal (e.g., Burgess & Hitch, 1992, 1996, 1999; Henson, 1998), is the assumption that the activation level of nodes corresponding to individual list items is temporally graded, such that the most active item node at any one time is the item that was in that particular position during list presentation. Burgess and Hitch (1992) altered the context-timing signal so that adjacent list items were more highly correlated than items that were temporally well separated. Through the workings of the context-timing signal, this slight modification allowed the Burgess and Hitch (1992) model to account for the large number of paired *item* transposition errors that are observed in the STM literature (e.g., Bjork & Healy, 1974).

However, in Glasspool's (1995) model, a 'chunk' refers either to a word in the word memory system, or a phoneme in the phoneme memory system. Thus, this model

"…completely ignores the fact that nonwords are themselves chunks" (Glasspool, 1995; p.26), and instead treats nonwords as strings of individual phonemes. Hence, when the stimuli are nonwords (e.g., *dieg* and *hayb*), each individual phoneme node (i.e., individual words in the case of word lists) is temporally sequenced by the context-timing signal. Due to the workings of this signal, when a phoneme substitution error occurs, the majority of errors should be paired transposition errors. Thus, if presented with *dieg* and *hayb*, the second and third phonemes in the first nonword are just as likely to be transposed as the first and second phonemes in the second nonword (i.e., *dgie* or *ayhb*).

Within the linguistics domain, however, research findings question the idea that nonwords are individual phoneme strings. For instance, the most common type of error that occurs when the experimental stimuli are nonwords is when phonemes from different syllables recombine to form a new or novel syllable (Ellis, 1980). These types of recombination errors are highly constrained in that vowels tend to substitute for other vowels (Treiman & Danis, 1988). Also, initial consonants tend to substitute for initial consonants and final consonants with final consonants (Ellis, 1980). Thus, maintaining the syllable structure of the list item (Treiman & Danis, 1988). Further, when an erroneous phoneme replaces the correct phoneme in a list, it generally shares articulatory features with the phoneme that it replaced (i.e., *F* with *S* or *B* with *P*; Conrad, 1962, 1964). Therefore, this model does not place linguistic constraints on phoneme ordering, and as such is unable to account for research findings that suggest that language specific constraints influence the types of errors that occur when participants are recalling a list of items, regardless of whether the stimuli are words or nonwords.

.

8.2.2. Current research findings in relation to Glasspool's (1995) short-term memory model

In contrast to a number of STM models (e.g., Brown, et al., 2000; Burgess & Hitch, 1992, 1999), Glasspool's (1995) model of STM treats words and nonwords in a similar way. Hence, this model predicts that, although performance should be lower for nonwords as compared to words, due to the additional support that words receive from lexical information that is assumed to be stored in LTM, the effect that phonemic similarity has on STM performance should be similar, regardless of stimulus type (i.e., words or nonwords). However, this model, as with the majority of STM models (e.g., Brown, et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1988, 1990a, 2002), is based on the distinctiveness assumption. Therefore, according to this model, as similarity increases order memory should decrease. Hence, as with the feature model of immediate memory (Nairne, 1988, 1990a, 2002), this model would need to be modified to reflect the idea that the speech trace for vowels is longer in duration and more acoustically intense than it is for consonants. However, unlike the feature model (Nairne, 1988; 1990a) this modification would necessitate the inclusion of an additional assumption, that of salience (i.e., or in linguistic terms, sonority).

In terms of the effect that phonemic similarity has on the recall of item information, it is unclear as to how Glasspool's (1995) model could provide a plausible explanation for the item recall advantage observed in the current thesis when the stimulus lists rhyme. For instance, two explanations have been proposed for the item recall advantage observed for rhyming lists of items: category cueing (Nairne, 2002) and sub-syllabic structure (Hartley & Houghton, 1996). However, Glasspool's (1995; see also, Brown et al., 2000; Burgess & Hitch, 1992, 1999) STM model does not incorporate an additional assumption with respect to the salience of list cues and the influence that this has on the recall of item information. Further, although this model was designed to serially order the phonemes within words, it does not place language specific constraints (i.e., in the English language the /mt/ never occur together in the onset of a word) on STM performance, nor does it provide a mechanism that is responsible for maintaining the syllable structure of presented items (i.e., onset-rhyme). As such, Glasspool's (1995) STM model does not provide the mechanisms necessary to account for the item recall advantaged for rhyming lists of items that was observed in the current thesis.

8.3. Where to next?

In short, slight modifications to existing STM models (e.g., Glasspool, 1995; Nairne, 1988, 1990a, 2002) cannot provide the mechanisms necessary to account for the current research findings. Therefore, if STM models are to successfully account for the current research findings it is important that these models not only include linguistic mechanisms that operate at the sub-syllabic level to constrain STM performance, but these mechanisms need to be consistent with current research findings within the linguistic domain.

8.4. Providing STM models with the mechanisms necessary to account for the current research findings

Hartley and Houghton (1996) have developed a ‘linguistically constrained model of STM for nonwords’ that treats vowels differently to consonants. Hartley and Houghton’s (1996) linguistically constrained model of STM for nonwords was originally based on Dell’s (1986, 1988) model of speech production and as such has much in common with this model. Although the Hartley and Houghton (1996) model has already been integrated into an existing verbal STM model (see Gupta & MacWhinney, 1997), Gupta and MacWhinney (1997) based this model on the original computational model of the phonological loop (i.e., Burgess & Hitch, 1992), and as such retains many of the flaws associated with this earlier model. Thus, if the current endeavour is to be successful (i.e., incorporating a linguistically constrained model into a current STM model), it will also be important to incorporate the constraints into a model “…that is capable of explaining a wider range of psychological data on verbal STM than any other current model of this type” (Burgess & Hitch, 1999; p. 577). As Burgess and Hitch (1999) suggest, a model such as Hartley and Houghton’s (1996) linguistically constrained model of STM for nonwords could be incorporated into their model. Currently, however, this modification has not been implemented. Hence, an illustration of how the Burgess and Hitch (1999) model could be modified to include linguistic constraints on STM performance is outlined below.

8.4.1. A modification to the Burgess and Hitch (1999) short-term memory model

The core components of the modified model are the same as in the original Burgess and Hitch (1999) model. For instance, the modified model consists of separate nodes to represent the context-timing signal, an item (i.e., syllables) level, input phonemes and output phonemes. As with the Burgess and Hitch (1999) model, the connections between the phoneme (i.e., output and input) and item layer are used to represent phonemic content. As such, Hartley and Houghton's (1996) idea of a content pathway is represented in this model by the connections between the item and phoneme layers. Thus, during list presentation, phonemic input activates the item representations and during recall, item representations activate the speech output layer. The components that have either been modified or added to the Burgess and Hitch (1999) model are depicted in grey (see Figure 8.3).

Please see print copy for Figure 8.3

Figure 8.3 Modified version of the architecture for the Burgess and Hitch (1999) model.

As with the original Burgess and Hitch (1999) model, there is a connection between the context-timing signal and the item layer. This connection is responsible for temporally sequencing presented list items. However, unlike the Burgess and Hitch (1999) model, the proposed model includes a connection between the context-timing signal and the input phoneme nodes. This connection is responsible for temporally sequencing the phonemes within a presented item. The connection between the context-timing and input phoneme layer is assumed to operate in a similar way to the connection between the context-timing and item layer. Thus, the pattern of activation is assumed to reset after every pause. Therefore, the same mechanism that is responsible for temporal grouping effects can be used to temporally separate groups of phonemes (i.e., one phoneme from another).

As has been suggested previously (i.e., refer section 2.7.2), Hartley and Houghton's (1996) structural pathway is responsible for placing language specific constraints on phoneme ordering and maintaining syllable structure. This pathway has been incorporated into the current model between the input and output phoneme nodes, via a bi-directional connection between the syllable template and item/syllable nodes. To gain a thorough understanding of the workings of the modified model, the following sections will trace the pathway that list items travel from presentation to recall.

8.4.2. Presentation

When an item is presented auditorily to the model, the phoneme (phonemic components) and context layers (temporal ordering) are activated simultaneously. This activation is sent to the item/syllable level via three pathways: The context-item, input phoneme-item (content information) and input phoneme-template-item (structural information) pathways. The structural pathway is responsible for maintaining the syllable structure of the presented item as well as placing language specific constraints on phoneme ordering. As with the Burgess and Hitch (1999) model, the most active item is selected at the item/syllable level via the CQ mechanism (see sections 2.3.1.2 and 2.4.2). Once an item has been selected, it is suppressed by the CQ mechanism. As each item is presented to the model the context signal is updated (Burgess & Hitch, 1996).

However, when an item is presented visually, information enters the system at the item/syllable level. This information is fed to the output phoneme layer via the content pathway (phonemic information) and the structural pathway. The incorporation of syllable structure into this pathway allows visual information entering the system via the item/syllable node, to activate input phoneme nodes by way of sub-vocalization. Thus, it allows language specific constraints to be placed on subvocal rehearsal and more generally (i.e., regardless of presentation modality), to reactivate input phoneme nodes.

8.4.3. Recall

At recall, the context-timing signal is reset to reflect the pattern that occurred when the first item was presented to the model (Burgess & Hitch, 1996). This activation spreads to the item/syllable nodes to select the most active node. The selected item/syllable node simultaneously feeds its activity to the phoneme layer via the item-phoneme connection (phonemic information) and the item-template-phoneme connection (structural information) and back again via these two connections. Hence, the most strongly activated item is then selected for output. As with the Burgess and Hitch (1999) model, the activation of the context-timing signal is sequential which makes the influence of serial order separable from the influence that similarity has on recall performance. However, in the modified model, the activation of the item-phoneme and item-template-phoneme nodes is parallel. Hence, allowing the phonemic content of list items to be influenced by the structural properties of an item and vice versa.

8.4.4. Limitations of the proposed model

Three points are worth noting about the proposed model. Firstly, although the proposed model builds on existing work, it has not at present been implemented. Hence, future work should aim to simulate the data reported in the current thesis. Secondly, in the Burgess and Hitch (1999) model, the item layer was used to refer to items of any length (i.e., one syllable or multi-syllabic items). The stimuli used in the current thesis were all one-syllable items. Thus, the term *item* in the proposed model refers to one-syllable words or nonwords. Hence, for the current work, the omission of a layer in which multi-syllabic items are represented is not problematic. However, the proposed model would be unable to provide an explanation for the effect that similarity has on STM performance for multi-syllabic items. Finally, in line with the Hartley and Houghton (1996) model, the proposed model includes a rhyme node. Although the data obtained in the current thesis support the inclusion of a rhyme node to represent syllable structure, it is not entirely clear whether the proposed STM model needs a specific rhyme node. For example, according to Church and Broadbent (1990) sets of oscillators deviate around modal frequencies. It is this variability in the frequencies of these oscillators that allows them to capture the external rhythm of the lists of items. The modal frequencies of oscillators would vary depending on the prosody of the list items, and as such may provide a biologically plausible mechanism in which to account for these differences without the need for a 'special' rhyme node. Although beyond the scope of the current thesis, future research should investigate whether the frequencies of temporal oscillators could provide a biologically plausible mechanism to explain the item recall advantage observed for rhyming lists of stimuli.

8.5. Qualifications of the present research and suggestions for future research

As with all research endeavours, there are a number of key decisions that were made throughout the current thesis that need to be addressed. Traditionally, experimental design justifications would be presented in an introduction to a particular experimental chapter. Although this is true for the most part, each of the experimental chapters in the current thesis is written in paper format. Hence, the discussion of five design issues that have arisen from the current work, and the implications of these decisions were deferred so that each concern could be addressed in greater detail.

One design issue that needs to be address is the decision to use a fixed list as compared to the memory span procedure. An advantage of using the memory span procedure in comparison to the fixed list length procedure is that it takes into account individual differences in STM capacity. However, a major criticism levelled at research into the PSE, and STM research more generally, stems from the idea that differences in the stimulus sets used across conditions influences the results obtained. For example, in comparison to the phonemically similar conditions, Fallon et al. (1999) used a different set of stimuli to construct their phonemically dissimilar lists. This same type of criticism cannot be levelled at the current research, in that the strength of the conclusions drawn from the current experiments stems from the way in which the stimulus lists were manipulated to form the different conditions across each experiment. For instance, the same stimulus set was used to construct the lists across all of the conditions for any one experiment. Thus, any performance differences observed across the different conditions in any one experiment in the current thesis (i.e., phonemically similar, dissimilar or lists that shared two phonemes), cannot be due to stimulus differences. This is important, in that if the stimuli were not identical across experimental conditions, any one of a number of variables could have influenced the results observed (e.g., word or biphone frequency, vowel length, number of liquids or stops in the initial as opposed to final position in a list). Further, if a memory span procedure was used to assess STM performance, different levels of performance across conditions would have meant that participants were presented with different items across conditions. This would have substantially weakened the conclusions drawn from the results observed in the current research. This is not to suggest that the memory span procedure should not be used to assess STM capacity when the experimental stimuli are phonemically similar. Rather, it was important that the current studies demonstrate that differential results are found in the literature depending on how similarity is operationally defined, before further studies into the effect that similarity has on STM performance are conducted.

To avoid the stimulus differences argument, and to make performance comparable to when the task was serial recall, the serial recognition task was designed such that a fixed list procedure was used. When the task is serial recall, three performance measures were obtained (correct-in-position, item recall and order accuracy). Item recall is a measure of overall recall ability, regardless of position. In contrast, the serial recognition task is a measure of order recognition. Further, the order accuracy measure obtained when the task is serial recall is a measure of order memory that takes into account individual differences in item recall ability. In contrast, the correct-in-position measure, as with the fixed length procedure when the task is serial recognition, is a measure of order memory that does not take into account these individual differences in overall item recall ability. Hence, the decision to compare the findings from the serial recognition task with those obtained when the task was serial recall and performance was measured using the correct-in-position criterion.

A third issue that needs to be addressed was the decision to increase the number of participants used when the task was serial recognition as compared to recall. One difference between these two tasks is that serial recall is a more sensitive measure of performance than the serial recognition task. For instance, on a serial recall task, when lists consist of six items, performance on each list is scored out of six. However, for the serial recognition task, there is only one data point per list (i.e., participants are either correct or incorrect). Thus, the serial recognition task is not as sensitive a measure of performance as the serial recall task (Gathercole et al., 2001). Hence, increasing the number of participants that performed the serial recognition as compared to the serial recall task was done in an attempt to make the findings observed across the serial recognition experiments more sensitive to the experimental manipulations.

Using a serial recognition task, Lian et al. (2001) found similar levels of performance, regardless of whether the stimuli were words or nonwords. Therefore, the decision to increase the number of items presented to participants from five to six when the task was the serial recognition of nonwords was based on previous research findings. However, as suggested previously, this task is not as sensitive to experimental manipulations as is the serial recall task. Also, the fixed length procedure does not take into account individual differences in item recall ability. This decision meant that a larger number of participants were responding randomly (as the task was potentially too difficult), which further decreased the sensitivity of this task to the experimental manipulations. In an attempt to combat this, the number of participants required to enhance the sensitivity of the task to the experimental manipulation was further increased for the serial recognition task when the stimuli were nonwords. Thus, although not an ideal situation, the appropriate design decisions have been made to compensate for this lack of sensitivity.

Finally, the decision was made to perform limited analyses on the types of errors that occurred when participants were performing the serial recall task. As has been suggested previously, the errors that occur when participants are recalling nonwords as compared to words are different (Ellis, 1980). For instance, although whole item substitution errors occur regardless of whether the stimuli are words or nonwords

(Hartley & Houghton, 1996), phoneme substitutions are more common when the stimuli are nonwords (Ellis, 1980). These types of findings could be used to strengthen the argument mounted that words and nonwords are processed differently (Lian et al., 2001). However, the current research suggests that once the demands that overt speech production have on STM performance are removed, the influence that similarity has on STM performance is similar for words and nonwords.

Further, as Hulme and Snowling (1992; see also Snowling et al., 1986) suggest, research to date has failed to separate phonological processing components which include phonological input and speech output from "…the phonological memory component (i.e., maintaining the phonological representation in the phonological loop)" (Bowey, 1997; p.298). This stems from the lack of an available STM measurement instrument that separates the memory component from other processing components. This is not to suggest that investigations into the types of errors that occur when the task is serial recall are fruitless. On the contrary, these types of studies have provided researchers with a rich source of information about different STM mechanisms and the level at which these mechanisms may operate (e.g., Conrad, 1962, 1964; Wickelgren, 1965a, 1965b, 1965c, 1966a, 1966b). However, if STM research is to move forward, researchers need to, not only acknowledge the influence that other processing components have on STM performance across different tasks, but design purer tests for particular processing components.

8.5.1. Modifying the serial recognition task

As has been suggested previously, all STM tasks have both benefits and limitations (e.g., Gathercole et al., 2001; Neath, 1997). For instance, one benefit to using the serial recall as compared to the serial recognition task is that the serial recall task is a more sensitive measure of experimental manipulations (Gathercole et al., 2001). In contrast, the serial recognition task, unlike serial recall, does not require participants to overtly articulate the presented list items (Martin et al., 1999). Thus, the results observed when using the serial recognition task are free from the influence that overt speech production may have on STM performance. However, because participants do not have to verbally recall the presented list items, a major problem with the task is that it is unclear whether participants are encoding the entire item or part of a presented item. For instance, a participant may be presented with the list *dog, hat, seam* and *fun* and then presented with the list *dog, seam, hat* and *fun*. To make a correct response (i.e., different), a participant may encode all of the item information or alternatively, part of the item (i.e., the initial consonant). This strategy may work effectively for either phonemically dissimilar or rhyming lists of items when each item in a list has a distinct initial phoneme. However, when the stimuli in a list are phonemically similar (i.e., *cat, cap, sat* and *gap* as compared with *cap, cat, sat* and *gap*), this type of strategy would not be as effective. Therefore, when the task is serial recognition, it is unclear as to what a participant is actually encoding.

Nevertheless there is a good reason to persevere with the serial recognition task: it provides a measure that separates the processing components of STM from a participant's speech production abilities. This is important in that Gathercole et al. (2001) have recently suggested that the findings observed from studies that use the serial recall task may be influenced by a participant's articulatory ability, especially when the experimental stimuli are nonwords. One way to avoid the issue of whether participants are encoding all, or part of a presented item would be to ask the participants about any strategies that they may have used to make the recognition judgments and omit the results of those participants who were using a strategy such as 'remember the initial consonant'. This is not advantageous, however, as participants may not admit to using this type of strategy and if they do use this strategy, the task is no longer measuring what it purports to measure.

Alternatively, a slight modification to the instructions given to participants may address this issue. For instance, currently, participants are required to say 'different' when they believe that the items on the second presentation have been presented in a different order. However, it may be better to ask participants to say the name of one of the items that they believe has been presented in a different order. In order for participants to make a correct response, they would need to encode all of the item information that was presented to them. Hence, this slight modification to the instructions given to participants when performing the serial recognition task would force participants away from a partial encoding strategy (i.e., remember the initial consonant). Thus, researchers could be confident that participants were attending to, not only the order in which items occurred but also encoding all of the item information presented. Although this type of modification would reintroduce articulation as a factor, the influence of articulation on STM performance would be substantially reduced.

8.5.2. Experiments designed to test the idea of a syllable frame

One of the implications for STM models identified in the current body of work is the need for these models to incorporate the linguistic mechanisms necessary to deal with the current research findings. Although linguistic models suggest that there is both a content and structural pathway, these types of models differ as to whether they are part of a single system (e.g., Dell, 1986, 1988) or separate pathways (Hartley & Houghton, 1996). This is one question in which future research should aim to address. A related research question that should also be addressed is the effect that syllable structure (i.e., the structural pathway) has on STM performance. Currently, research into the effect that the syllable structure of experimental items has on STM performance for phonemically similar lists of items has not been investigated.

There are a number of ways in which future research could investigate the effect that syllable structure has on STM performance. For instance, lists of items could be constructed that vary the syllable structure of the list items. Using the serial recall task, performance across five-item lists of words that consist of CVC syllables (i.e., 15 phonemes) that shared a common vowel could be compared with lists that consist of a mixture of CV, VC, CVC, CCVC and CVCC items (i.e., 15 phonemes). Thus, keeping both the number of phonemes and the vowel constant across lists. Performance across these types of lists could be contrasted with performance on lists of phonemically dissimilar items.

An effect of syllable structure manipulation would lend support to the idea that syllable structure influences STM performance (Dell, 1986, 1988; Hartley & Houghton, 1996; Seveld & Dell, 1984; Seveld, Dell & Cole, 1995). However, if no performance differences were found when syllable structure was manipulated (i.e., either phonemically similar or phonemically dissimilar lists), this would lend support to the claim that syllables are *chunks,* and as such if the number of phonemes is held constant across lists (i.e., 15 phonemes in each list), syllable structure should not influence STM performance (Dell et al., 1993; Estes, 1972; Johnson, 1972).

8.6. Summary and conclusions

For decades, the PSE has been described as the finding that although the same number of items are recalled regardless of whether the stimuli are phonemically similar or distinct (Wickelgren, 1965d), phonological similarity impairs our ability to recall a list of items in the correct order (e.g., Baddeley et al., 1984; Conrad et al., 1966; Conrad & Hull, 1964; Cowan et al., 1991; Li et al., 2000). Recent research, however, has questioned the stability of the PSE on STM performance. For instance, although the standard detrimental effect of similarity for order memory is observed (e.g., Farrell & Lewandowsky, 2003), when the recall of item information is measured, some studies have found no difference in item recall levels between phonemically similar and dissimilar lists (Poirier & Saint-Aubin, 1996), whereas others have found that phonemic similarity can facilitate (Fallon et al., 1999) or have a detrimental effect (Coltheart, 1993) on the recall of item information. Hence, to gain a thorough understanding of the influence that phonological similarity has on STM performance, it is important that researchers examine the effect that similarity has on both order memory and the recall of item information.

At present there are two competing explanations that are derived from extant STM model to account for the effect that phonological similarity has on order memory. According to non-linguistic models of STM (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a), order memory is influenced by the distinctiveness of list items in relation to the other items in a list. Hence, according to these types of models, as similarity increases order memory should decrease. In contrast, psycholinguistic models of STM (e.g., Gupta & MacWhinney, 1997; Hartley & Houghton, 1996) suggest that although sharing any form of similarity should have a detrimental effect on order memory, when the overlapping phoneme is the vowel, a further order memory impairment should be observed.

In terms of the effect that phonological similarity had on order memory, the current research found that when similarity was held constant (equal with respect to the number of phonemically similar phonemes that list items overlapped on), regardless of stimuli type (i.e., words vs. nonwords) or the task performed (i.e., serial recall vs. serial recognition) order memory was better for C_C as compared either CV_ or _VC lists of items. Hence, models of STM that are based on the ratio rule (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a), cannot adequately explain the effect that similarity has on order memory. Rather, the findings are more consistent with Hartley and Houghton's (1996) linguistically constrained model of STM for nonwords that assumes that STM performance is influenced by linguistic mechanisms, such as sonority, that operate at the sub-syllabic level.

Both psycholinguistic, as well as a number of non-linguistic models of STM also provide explanations to account for the item recall advantage observed in the research literature when the experimental stimuli rhyme (e.g., Gathercole et al., 1982; Fallon et al., 1999). For instance, according to newer versions of the feature model, retrieval cues such as a shared rhyme ending, can be used to limit the size of the secondary memory search set and hence, facilitate the recall of item information (Nairne, 2002; Nairne & Kelley, 1999; see also Tehan & Fallon, 1999 for a similar argument). Thus, according to these models, a beneficial effect of similarity for the recall of item information should be observed whenever item similarity can be used as a cue to limit the size of the secondary memory search set (i.e., _VC, CV_ or C_C lists). In contrast, Hartley and Houghton's (1996) model of STM for nonwords, is based on linguistic research that suggests that syllables are divided into an onset, or initial phoneme or phoneme cluster and a rhyme (Treiman, 1983). Accordingly then, item recall should benefit when list items share a common rhyme ending. However, when list items do not rhyme (i.e., CV_ or C_C components), this item recall advantage should be absent (or at least minimal).

In terms of the effect that phonological similarity had on the recall of item information, when the stimuli were words, the current research findings were consistent with predictions based on the Hartley and Houghton (1996) model, in that an item recall advantage was observed for rhyming lists of items, and was absent (or at least minimal) when the stimulus lists shared either the CV_ or C_C components. However, when the stimuli were nonwords, the findings were somewhat mixed. For instance, an item recall advantage was observed not only for rhyming lists of items, but also when the stimulus lists shared the C_C component. One suggestion proposed to account for these discrepant findings was that in comparison to when item information is measured using a free recall task, which does not place order constraints on recall performance (e.g., Bjork & Whitten, 1974; Glanzer & Cunitz, 1966), the results observed in the current thesis may have underestimated the true extent to which similarity influences the recall of item information. Thus, future research may need to investigate the effect that similarity has on the recall of item information with a purer measure of item recall. However, when taken together, the current findings do lend support to the suggestion that there is a beneficial effect on the recall of item information that is dependent on how phonemic similarity is operationally defined.

Currently a debate rages as to whether the same (Fallon et al., in press) or different (Lian et al., 2001) processes are involved in word and nonword recall. The findings from the current study can help illuminate this issue in that, almost identical patterns of results across word and nonword lists were found when performance was measured using the order accuracy criterion. Hence, these findings strengthen the claim that words and nonwords are processed in a similar way (Fallon et al., in press). Although, when performance was measured using the correct-in-position criterion, order memory was worse for phonemically dissimilar as compared to nearly all of the phonemically similar nonword conditions. This can be contrasted with the findings from the word studies that suggest that correct-in-position performance is higher for phonemically dissimilar lists, regardless of how similarity is operationally defined.

One suggestion proposed to account for these discrepant findings was that articulatory skill influences recall performance, especially when the stimuli are nonwords (Gathercole et al., 2001). When performance was measured using a task that is not influenced by overt speech production constraints (i.e., serial recognition) an identical pattern of results was observed, regardless of stimulus type (i.e., words vs. nonwords). Further, the patterns of results were identical to those obtained for words using the correct-in-position measure of performance. These findings not only suggest that words and nonwords are processed in a similar way (Fallon et al., in press), but that the speech output component of phonological processing influences the results obtained on serial recall tasks to a greater extent when the experimental stimuli are nonwords. Finally, the results suggest that once the demands that overt speech production have on STM performance are removed (i.e., using a serial recognition task), the influence that similarity has on STM performance at the sub-syllabic level persists. Hence, the current demonstration of a sub-syllabic influence on STM performance that is independent of speech production processes suggests that this is a robust STM effect.

In summary, a number of clear conclusions can be drawn from the current research findings. Firstly, the findings from the current thesis suggest that regardless of whether the stimuli are words or nonwords, STM performance is influenced by linguistic mechanisms that operate at the sub-syllabic level. This suggests that similar sub-syllabic processes (i.e., syllable structure and sonority) influence STM performance when recalling either words or nonwords. Secondly, that the influence that phonemic similarity has on STM performance remains once the demands that overt speech production have on STM performance are removed. Finally, the findings from the current thesis suggest that it is no longer sufficient for STM models to use a 'phonemic overlap' or distinctiveness' argument to account for the effect that phonemic similarity has on order memory. Rather, if STM models are to successfully provide an explanation for the effect that phonemic similarity has on STM performance, they need to incorporate linguistic constraints, such as syllable structure and sonority, into existing STM model.

The goal for STM researchers is to construct experiments that challenge the mechanisms that have been outlined by modellers to explain certain effects. In contrast, the challenge that STM modellers face is to provide STM models with the mechanisms that are capable of dealing with these findings. As suggested previously, the majority of STM models rely on the distinctiveness assumption to provide an explanation for the effect that similarity has on order performance (e.g., Brown et al., 2000; Burgess & Hitch, 1992, 1999; Nairne, 1990a). However, the current research findings challenge this view. Further, the findings from the current thesis also suggest that slight modifications to current STM models (e.g., Glasspool, 1995; Nairne, 1988, 1990a, 2002), will not adequately deal with the current research findings. Hence, to provide an adequate explanation for the current research findings, STM models need to incorporate mechanisms that are based on linguistic research. Further, the integration of Hartley and Houghton's (1996) linguistically constrained model with the Burgess and Hitch (1999) model, which is arguable the most advanced computational model of verbal STM has been proposed. As such, the value of the current body of research lies in the sizable contribution it makes to further our understanding of the effect that similarity has on STM performance, and in the implications that these findings have for STM models.

The last few years have seen a change in focus from that of delineating effects to providing biologically plausible mechanisms to explain these effects. This has provided cognitive psychology with the tools necessary to solve problems such as the serial ordering of behaviour, that have plagued STM models for over fifty years (e.g., Lashley, 1951). This is an exciting time within the cognitive sciences, and especially, within the field of cognition. The major contribution that the current body of work bestows to cognitive psychology, although substantial, may not be found in the questions that the current research has answered. Rather, its importance may lie in the research questions that proliferate from this body of work. Only time will provide an answer to this question.

References

Andrade, J. (2001). An introduction to working memory. In J. Andrade (Ed.), *Working memory in perspective* (pp. 3-30). Hove, East Sussex: Psychological Press Ltd.

Ashcraft, M. H. (2002). *Cognition* (3rd ed.). Upper Saddle River, New Jersey: Prentice Hall.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control processes of short-term memory. *Scientific American, 225*, 82-90.

Avons, S. E., Wright, K. L., & Pammer, K. (1994). The word-length effect in probed and serial recall. *The Quarterly Journal of Experimental Psychology, 47A*, 207-231.

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.

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology Human Experimental Psychology, 20A*, 249-264.

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

Baddeley, A. D. (1997). *Human memory: Theory and practice* (Revised ed.). East Sussex, UK: Psychological Press.

Baddeley, A. D., Chincotta, D., Stafford, L., & Turk, D. (2002). Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *The Quarterly Journal of Experimental Psychology, 55A*, 353-369.

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology, 36A*, 233-252.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multi-component model. In A. Miyake & P. Shah (Eds.), *Model of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge, United Kingdom: Cambridge University Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589.

Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin, 74*, 81-99.

Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology, 36*, 701-711.

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior, 13*, 80-97.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6*, 173-189.

Bowey, J. A. (1997). What does nonword repetition measure? A reply to Gathercole and Baddeley. *Journal of Experimental Child Psychology, 67*, 295-301.

Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35*, 345-367.

Broadbent, D. E., & Bradbent, M. H. (1981). Recency effects in visual memory. *Quarterly Journal of Experimental Psychology, 33A*, 1-15.

Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed? *Journal of Memory and Language, 34*, 594-621.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107*, 127-181.

Burgess, N. (1995). A solvable connectionist model of immediate recall of ordered lists. *Advances in Neural Information Processing, 7*, 51-58.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31*, 429-460.

Burgess, N., & Hitch, G. J. (1996). A connectionist model of STM for serial order. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 51-72). Hove, England: Psychological Press.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.

Campbell, R., & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in highly literate subjects: A developmental case with associated deficits of phonemic processing and awareness. *Quarterly Journal of Experimental Psychology, 37A*, 435-476.

Church, R. M., & Broadbent, H. A. (1990). Alternative representation of time, number, and rate. *Cognition, 37*, 55-81.

Church, R. M., & Broadbent, H. A. (1992). A connectionist model of timing. In Commons, M.L., S. Grossberg & J. E. R. Staddon (Eds.), *Neural network models of conditioning and action* (pp. 225-240). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Clegg, F., & Warrington, E. K. (1994). Four easy memory tests for older adults. *Memory, 2*, 167-182.

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

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term memory recall of repeated and novel word lists. *Memory and Cognition, 21*, 539-545.

Conrad, R. (1960). Serial order intrusions in immediate memory. *British Journal of Psychology, 51*, 45-48.

Conrad, R. (1962). An association between memory errors and errors due to acoustic masking of speech. *Nature, 193*, 1314-1315.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Experimental Psychology, 33A*, 1-15.

Conrad, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning and Verbal Behavior, 4*, 161-169.

Conrad, R. (1967). Interference or decay over short retention intervals? *Journal of Verbal Learning and Verbal Behavior, 6*, 49-54.

Conrad, R., Baddeley, A. D., & Hull, A. J. (1966). Rate of presentation and the acoustic similarity effect in short-term memory. *Psychonomic Science, 5*, 233-234.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429-432.

Conrad, R., & Hull, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science, 10*, 135-136.

Corballis, M. C. (1966). Rehearsal and decay in immediate recall of visually and aurally presented items. *Canadian Journal of Psychology, 20*, 43-51.

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language, 31*, 668-684.

Cowan, N., & Kail, R. (1996). Covert processes and their development in short-term memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 29-50). East Sussex, England: Psychological Press.

Cowan, N., Keller, T. A., Hulme, C., Roodenrys, S., McDougall, S., & Rack, J. (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory and Language, 33*, 234-250.

Cowan, N., Saults, J. S., Winterowd, C., & Sherk, M. (1991). Enhancement of 4-year-old children's memory span for phonologically similar and dissimilar word lists. *Journal of Experimental Child Psychology, 51*, 30-52.

Cowan, N., Wood, N. L., Wood, P. K., Keller, T. A., Nugent, L. D., & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127*, 141-160.

Craik, F. I. M. (1969). Modality effects in short-term storage. *Journal of Verbal Learning and Verbal Behavior, 8*, 658-664.

Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior, 9*, 143-148.

Crowder, R. G. (1972). Visual and auditory memory. In J. F. Kavanagh & I. G. Mallingly (Eds.), *Language by ear and by eye* (pp. 251-276). Cambridge, MA: MIT Press.

Crowder, R. G. (1979). Similarity and order in memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 13, pp. 319-353). New York: Academic Press.

Dale, H. C. A. (1964). Retroactive interference in short-term memory. *Nature, 203*, 1408.

Dale, H. C. A., & Gregory, M. (1966). Evidence of semantic coding in short-term memory. *Psychonomic Science, 5*, 75-76.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language, 27*, 124-142.

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17*, 149-195.

Drewnowski, A. (1980). Memory functions for vowels and consonants: A reinterpretation of acoustic similarity effects. *Journal of Verbal Learning and Verbal Behavior, 19*, 176-193.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. New York: Dover. (Original work published in 1890).

Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior, 19*, 624-634.

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. *American Scientist, 70*, 607-615.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161-190). Washington DC: Winston.

Estes, W. K. (1985). Memory for temporal information. In J. A. Michin & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 151-168). Berlin, Germany: Springer-Verlag.

Estes, W. K. (1997). Processes of memory loss, recovery and distortion. *Psychological Review, 104*, 148-169.

Fallon, A. B., Groves, K., & Tehan, G. (1999). Phonological similarity and trace degradation in the serial recall task: When CAT helps RAT, but not MAN. *International Journal of Psychology, 34*, 301-307.

Fallon, A. B., Mak, E., Tehan, G., & Daly, C. (in press). Lexicality and phonological similarity in serial recall: A challenge for the retrieval-based model. *Memory*.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin and Review, 9*, 59-79.

Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 838-849.

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24*, 200-209.

Fudge, E. C. (1969). Syllables. *Journal of Linguistics, 5*, 253-286.

Galotti, K. M. (1999). *Cognitive psychology* (2nd ed.). Belmont, CA: Wadsworth Publishing Company.

Gathercole, S. E. (1995). Nonword repetition: More than just a phonological output task. *Cognitive Neuropsychology, 12*, 857-861.

Gathercole, S. E. (1997). Models of verbal short-term memory. In M. A. Conway (Ed.), *Cognitive models of memory* (pp. 13-45). Cambridge, MA: MIT Press.

Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of nonword repetition*. UK: Psychological Corporation.

Gathercole, S. E., Gardiner, J. M., & Gregg, V. H. (1982). Modality and phonological similarity effects in serial recall: Does one's own voice play a role? *Memory and Cognition, 10*, 176-180.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33*, 966-979.

Gathercole, S. E., & Martin, A. J. (1996). Interactive processes in phonological memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 73-100). East Sussex, UK: Psychological Press.

Gathercole, S. E., & McCarthy, R. A. (1994). Memory tools for the researcher and clinician. *Memory, 2*, 97-101.

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology, 54A*, 1-30.

Gathercole, S. E., Service, E., Hitch, G., Adams, A., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*, 65-77.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991a). Nonword repetition, phonological memory, and vocabulary: A reply to Snowling, Chiat, and Hulme. *Applied Psycholinguistics, 12*, 375-379.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991b). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12*, 349-367.

Gillund, G., & Shriffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1-67.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351-360.

Glasspool, D. W. (1995). Competitive queuing and the articulatory loop. In J. Levy, D. Bairaktaris, J. Bullinaria & P. Cairns (Eds.), *Connectionist models of memory and language* (pp. 5-30). London: UCL Press.

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders*. Philadelphia: Lea and Febiger.

Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Guildford, J. P., & Dallenbach, K. M. (1925). The determination of memory span by the method of constant stimuli. *American Journal of Psychology, 36*, 621-628.

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

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language, 35*, 1-31.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 644-655.

Healy, A. F., Fendrich, D. W., Cunningham, T. F., & Till, R. E. (1987). Effects of cuing on short-term retention of order information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 413-425.

Hebb, D. O. (1961). Brain mechanisms and learning. In J. F. Delafresnaye (Ed.), *Distinctive features of learning in the higher animal* (pp. 37-46). London: Oxford University Press.

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36*, 73-137.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *The Quarterly Journal of Experimental Psychology, 49A*, 80-115.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411-428.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly Journal of Experimental Psychology, 49A*, 116-139.

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

Houghton, G., Glasspool, G. W., & Shallice, T. (1994). Spelling and serial recall: Insights from a competitive queuing model. In D. A. Brown & N. C. Ellis (Eds.), *Handbook of normal and distributed spelling* (pp. 365-404). Chichester, England: John Wiley.

Houghton, G., Hartley, T., & Glasspool, D. W. (1996). The representation of words and nonwords in short-term memory: Serial order and syllable structure. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 101-127). East Sussex, UK: Psychological Press.

Houghton, G. D. W., & Hartley, T. (1996). Parallel models of serial behaviour: Lashley revisited. *Psyche: An interdisciplinary Journal of Research of Consciousness, 2*, (25).

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

Hulme, C., Roodenrys, S., Brown, G. D. A., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of Psychology, 86*, 527-536.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 1217-1232.

Hulme, C., & Snowling, M. (1992). Deficits in output phonology: An explanation of reading failure? *Cognitive Neuropsychology, 9*, 47-72.

Huttenlocher, J., & Newcombe, N. (1976). Semantic effects on ordered recall. *Journal of Verbal Learning and Verbal Behavior, 15*, 387-399.

Jacobs, J. (1887). Experiments on "prehension". *Mind, 12*, 75-79.

Jahnke, J. C. (1969). The Ranschburg effect. *Psychological Review, 76*, 592-605.

James, W. (1890). *Principles of psychology*. New York: Holt.

Johnson, G. J. (1991). A distinctiveness model of serial learning. *Psychological Review, 98*, 204-217.

Johnson, N. F. (1972). Coding processes in human memory. In A. W. Melton & E. Martin (Eds.), *Organization and the concept of a memory code* (pp. 161-190). Washington, DC: Winston.

Kauffman, S. A. (1993). *The origins of order*. Oxford, England: Oxford University Press.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37*, 295-311.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-136). New York: Wiley.

Laughery, K. R., & Pinkus, A. L. (1966). Short-term memory: Effects of acoustic similarity, presentation rate and presentation mode. *Psychonomic Science, 6*, 285-286.

Lee, C. L. (1992). The perturbation model of short-term memory: A review and some further developments. In A. F. Healy, R. M. Kosslyn & R. M. Shiffrin (Eds.), *From learning theory to cognitive processes: Essays in honor of William K. Estes* (pp. 119-141). Hillsdale, NJ: Erlbaum.

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 149-169.

Lewandowsky, S., & Murdock, J., B.B. (1989). Memory for serial order. *Psychological Review, 96*, 25-57.

Li, X., Schweickert, R., & Gandour, J. (2000). The phonological similarity effect in immediate recall: Positions of shared phonemes. *Memory and Cognition, 28*, 1116-1125.

Lian, A., Karlsen, P. J., & Winsvold, B. (2001). A re-evaluation of the phonological similarity effect in adults' short-term memory of words and nonwords. *Memory, 9*, 281-299.

Lockhart, R. S., & Murdock, B. B. J. (1970). Memory and the theory of signal detection. *Psychological Bulletin, 74*, 100-109.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8*, 323-350.

Martin, R. C., & Breedin, S. D. (1992). Dissociations between speech perception and phonological short-term memory deficits. *Cognitive Neuropsychology, 9*, 509-534.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language, 41*, 3-29.

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General, 118*, 346-373.

McNichol, D. (1972). *A primer of signal detection theory*. London: Allen and Unwin.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior, 2*, 1-21.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*, 3-19.

Miikkulainen, R. (1990). A distributed feature map model of the lexicon. In M. Piattelli-Palmarini (Ed.), *The twelfth annual conference of the cognitive science society* (pp. 447-454). Hillsdale, NJ: Lawrence Erlbaum Associated, Inc.

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

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology, 10*, 465-501.

Morton, J., & Patterson, K. E. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. E. Patterson & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 91-118). London: Routledge.

Murdock Jr, B. B. (1960). The distinctiveness of stimuli. *Psychological Review, 67*, 16-31.

Murdock Jr, B. B. (1968). Serial order effects in short-term memory. *Journal of Experimental Psychology: Monograph Supplement, 76*, 1-15.

Murdock Jr, B. B. (1974). *Human memory: Theory and data*. Polomac, Md: Lawrence Erlbaum Associates.

Murdock Jr, B. B. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General, 105*, 191-216.

Murdock Jr, B. B. (1983). A distributed memory model for serial-order information. *Psychological Review, 90*, 316-338.

Murdock Jr, B. B., & Walker, K. D. (1969). Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior, 8*, 665-676.

Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing: Principals and applications* (4th ed.). Upper Saddle River, New Jersey: Prentice-Hall International, Inc.

Murray, A., & Jones, D. M. (2002). Articulatory complexity at item boundaries in serial recall: The case of Welsh and English digit span. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*, 594-598.

Murray, D. J. (1966). Vocalization-at-presentation and immediate recall with varying recall methods. *Quarterly Journal of Experimental Psychology, 18*, 9-18.

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*, 679-684.

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory and Cognition, 16*, 343-352.

Nairne, J. S. (1990a). A feature model of immediate memory. *Memory and Cognition, 18*, 251-269.

Nairne, J. S. (1990b). Similarity and long-term memory for order. *Journal of Memory and Language, 29*, 733-746.

Nairne, J. S. (1991). Positional uncertainty in long-term memory. *Memory and Cognition, 19*, 332-340.

Nairne, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science, 3*, 199-202.

Nairne, J. S. (2002a). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81.

Nairne, J. S., & Kelley, M. R. (1999). Reversing the phonological similarity effect. *Memory and Cognition, 27*, 45-53.

Nairne, J. S., Neath, I., Serra, M., & Byun, E. (1997). Positional distinctiveness and the ratio rule in free recall. *Journal of Memory and Language, 37*, 155-166.

Nairne, J. S., & Neumann, C. (1993). Enhancing effects of similarity on long-term memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 329-337.

Neath, I. (1993a). Contextual and distinctive processes and the serial position function. *Journal of Memory and Language, 32*, 820-840.

Neath, I. (1993b). Distinctiveness and serial position effects in recognition. *Memory and Cognition, 21*, 689-698.

Neath, I. (1997). Modality, concreteness, and set-size effects in a free reconstruction of order task. *Memory and Cognition, 25*, 256-263.

Neath, I. (1999). Modelling the disruptive effect of irrelevant speech on order information. *International Journal of Psychology, 34*, 416-424.

Neath, I., & Nairne, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin and Review, 2*, 429-441.

Nimmo, L. M., & Roodenrys, S. (2002). Syllable frequency effects on phonological short-term memory tasks. *Applied Psycholinguistics, 23*, 643-659.

Nimmo, L. M., & Roodenrys, S. (2004). Investigating the phonological similarity effect: Syllable structure and the position of common phonemes. *Journal of Memory and Language, 50*, 245-258.

Nimmo, L. M., & Roodenrys, S. (submitted). Investigating the phonological simialrity effect: Are nonwords an alien form? *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Nosofsky, R. M. (1986). Attention, similarity and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.

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

Pinel, J. P. J. (1997). *Biopsychology* (3rd ed.). Needham Heights, MA: Allyn and Bacon.

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology, 48A*, 384-404.

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, Word frequency, item identity and item position. *Canadian Journal of Experimental Psychology, 50*, 408-412.

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology, 17*, 132-138.

Roodenrys, S., Hulme, C., & Brown, G. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology, 56*, 431-442.

Roodenrys, S., & Quinlan, P. T. (2000). The effects of stimulus set size and word frequency on verbal serial recall. *Memory, 8*, 71-78.

Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology, 21*, 137-147.

Ryan, J. (1969b). Temporal grouping, rehearsal and short-term memory. *Quarterly Journal of Experimental Psychology, 21*, 148-155.

Saint-Aubin, J., & Poirier, M. (1999a). Semantic similarity and immediate serial recall: Is there a detrimental effect on order information? *The Quarterly Journal of Experimental Psychology, 52A*, 367-394.

Saint-Aubin, J., & Poirier, M. (1999b). The influence of long-term memory factors on immediate serial recall: An item and order analysis. *International Journal of Psychology, 34*, 347-352.

Saint-Aubin, J., & Poirier, M. (2000). Immediate serial recall of words and nonwords: Tests of the retrieval-based hypothesis. *Psychonomic Bulletin and Review, 7*, 332-340.

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

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory and Cognition, 21*, 168-175.

Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 419-425.

Schweickert, R., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science, 1*, 74-77.

Service, E. (1998). The effect of word length on immediate serial recall depends on phonological complexity, not articulatory duration. *The Quarterly Journal of Experimental Psychology, 51A*, 283-304.

Service, E., & Maury, S. (2003). All parts of an item are not equal: Effects of phonological redundancy on immediate recall. *Memory and Cognition, 31*, 273-284.

Service, E., Maury, S., & Luotoniemi, E. (in press). Forgetting and redintegration of consonants and vowels in pseudoword lists. *Memory*.

Seveld, C. A., Dell, G. S., & Cole, J. S. (1995). Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language, 34*, 807-820.

Shallice, T., & Butterworth, B. (1977). Short-term memory impairments and spontaneous speech. *Neuropsychologica, 15*, 729-735.

Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech-perception and production separate processes? *Quarterly Journal of Experimental Psychology, 37(A)*, 507-532.

Simon, H. A. (1974). How big is a chunk? *Science, 183*, 482-488.

Snowling, C., & Hulme, C. (1989). A longitudinal case study of developmental phonological dyslexia. *Cognitive Neuropsychology, 6*, 379-401.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics, 12*, 369-373.

Snowling, M., Goulandris, N., Bowlby, M., & Howel, P. (1986). Segmentation and speech perception in relation to reading skill: A developmental analysis. *Journal of Experimental Child Psychology, 41*, 489-507.

Sperling, G., & Speelman, R. G. (1970). Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of human memory* (pp. 151-202). New York: Academic Press.

Sternberg, R. J. (1999). *Cognitive psychology* (2nd ed.). Orlando, Florida: Harcourt Brace College Publishers.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*, 421-457.

Tehan, G., & Fallon, A. B. (1999). A connectionist model of short-term cued recall. In J. Wiles & T. Dartnell (Eds.), *Prospectives in cognitive science* (pp. 221-237). Stamford: Ablex.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15*, 49-74.

Treiman, R. (1984). On the status of final consonant clusters in English syllables. *Journal of Verbal Learning and Verbal Behavior, 23*, 343-356.

Treiman, R. (1986). The division between onsets and rhymes in English syllables. *Journal of Memory and Language, 25*, 476-491.

Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 145-152.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29*, 66-85.

Treisman, M., Cook, N., Naish, P. L. N., & MacCrone, J. K. (1994). The internal clock: Electroencephalographic evidence for oscillatory processes underlying time perception. *The Quarterly Journal of Experimental Psychology, 47A*, 241-289.

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behaviour, 23*, 151-161.

Vallar, G., & Shallice, T. (1990). *Neuropsychological impairments of short-term memory*. Cambridge, England: Cambridge University Press.

Watkins, M. J. (1977). The intricacy of memory span. *Memory and Cognition, 5*, 529-534.

Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior, 13*, 430-447.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 442-452.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72*, 89-104.

Wechsler, D. (1991). *Manual for the Wechsler intelligence scale for children (WISC-III)* (3rd ed.). San Antonio, Texas: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler adult intelligence scale (WAIS-III)* (3rd ed.). San Antonio, Texas: The Psychological Corporation.

Wells, B. (1995). Phonological considerations in repetition tests. *Cognitive Neuropsychology, 12*, 847-855.

Whiteman, H. L., Nairne, J. S., & Serra, M. (1994). Recognition and recall-like processes in the long-term reconstruction of order. *Memory, 2*, 275-294.

Wickelgren, W. A. (1965a). Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 4*, 53-61.

Wickelgren, W. A. (1965b). Distinctive features and errors in short-term memory for English vowels. *Journal of the Acoustical Society of America, 38*, 583-588.

Wickelgren, W. A. (1965c). Similarity and intrusions in short-term memory for consonant-vowel diagrams. *Quarterly Journal of Experimental Psychology, 17*, 241-246.

Wickelgren, W. A. (1965a). Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology, 70*, 102-108.

Wickelgren, W. A. (1965d). Short-term memory for phonologically similar lists. *American Journal of Psychology, 78*, 567-574.

Wickelgren, W. A. (1966). Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Society of America, 39*, 388-398.

Wickelgren, W. A. (1966b). Phonemic similarity and interference in short-term memory for single letters. *Journal of Experimental Psychology, 71*, 396-404.

Wickelgren, W. A. (1966c). Short-term recognition memory for single letters and phonemic similarity of retroactive interference. *Quarterly Journal of Experimental Psychology, 18*, 55-62.

Appendix A - Stimulus Lists used in Chapter 4 (Study One)

***Table A1**. Rhyming Word Set Used for Experiment 1*

name	maim	came	lame	dame	shame
wait	bait	mate	date	fate	rate
sail	mail	fail	pail	gail	tail
lace	face	mace	race	case	base
hide	ride	side	lied	tied	died
shine	mine	sign	line	wine	fine
bile	mile	tile	dial	rile	file
gnome	foam	roam	dome	comb	home
loan	bone	cone	phone	hone	sewn
coke	soak	choke	woke	poke	folk
goal	dole	bowl	foal	soul	pole
bag	wag	lag	rag	hag	tag
hang	rang	tang	gang	sang	pang
ram	sham	ham	lamb	mam	dam
cap	gap	rap	sap	chap	map
fin	win	sin	chin	pin	kin
whip	lip	kip	chip	tip	sip
dig	pig	big	wig	rig	fig
sing	ring	wing	ding	ting	ping
seize	tease	peas	knees	bees	cheese
bead	seed	heed	need	lead	weed
reap	seep	weep	sheep	beep	leap
sheik	meek	peek	cheek	teak	seek
sun	ton	run	gun	shun	done
rug	bug	hug	tug	mug	dug
tub	sub	cub	rub	pub	hub
putt	cut	gut	hut	rut	mutt
mock	wok	hock	rock	sock	knock
log	bog	cog	hog	dog	fog
lot	what	rot	not	pot	cot

Table A2. *CV_ Word Set used for Experiment 2*

mace	maim	main	mate	make	maze
sane	safe	sake	same	save	sail
wade	waif	whale	wain	wake	weighs
rail	rake	race	rain	rate	rave
live	life	light	lime	line	lice
rhyme	rice	rife	right	ripe	rise
ties	tight	time	type	tide	tile
while	white	wife	wise	wide	wine
bird	burn	burp	burg	burrs	birch
purred	pearl	perk	pert	purrs	perch
cab	can	cap	cat	cash	caff
hack	have	hash	hag	had	hang
man	mass	map	mate	mad	mag
sack	sad	sag	sang	sap	sat
lick	lid	limb	lip	lit	liv
pick	pig	pin	pit	pip	pill
bid	bill	bit	big	bin	bitch
seas	seal	seem	seek	seat	scene
wheeze	week	wean	wheel	weed	wheat
beak	bead	beam	beep	bees	beat
dung	dull	done	duck	duff	dumb
huff	hum	hush	hut	hub	hutch
mud	mull	mutt	muck	much	mush
rub	rum	rung	rush	rough	run
lob	lock	long	lop	loss	lot
cob	cod	con	cop	cot	cough
sob	sock	sod	song	sop	sot
poured	porch	pork	pawn	pause	port
ward	walk	wharf	warn	wart	wars
half	hard	harm	heart	halve	harsh

Table A3. *C_C Word Set Used for Experiment 3*

bide	beard	bod	bud	bored	bird
batch	birch	beach	bitch	botch	butch
bought	bet	boot	but	bat	bait
bile	bowl	bill	bull	bail	ball
choose	cheese	chose	chars	chores	cheers
fort	feet	fit	foot	fat	fate
fall	foul	feel	fool	full	fail
fees	fears	foes	furs	phase	fuzz
fawn	fern	phone	fun	fan	feign
guard	gored	guide	goad	geared	god
gut	got	get	git	goat	gate
ham	harm	home	him	whom	hum
hard	hide	heard	heed	hid	had
heart	hurt	heat	hut	hat	hoot
lurk	leak	look	lick	lake	lack
lawn	lane	line	learn	loan	lean
mess	mice	mace	mass	morse	moose
mourn	main	mine	moon	men	man
porch	patch	perch	poach	pitch	peach
pars	pays	purrs	piers	pause	pies
pal	pearl	pole	peel	pool	pail
rook	ruck	rake	reek	rack	wreck
rhyme	roam	room	rum	ream	ram
roared	raid	ride	road	red	read
wrote	route	rut	rat	rate	wrought
soars	size	sues	sears	seas	sews
suck	sick	sock	soak	sake	sack
sap	soap	sip	seep	sop	soup
walk	work	woke	wick	week	wack
warn	wian	wine	win	wean	one

Appendix B – ANOVA tables for the major statistical analyses performed for Chapter 4 (Study One)

Table B 1: Repeated measures 3 x 6 ANOVA on the mean number of correct responses using the correct-in-position measure for rhyming lists (Experiment 1).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	1259.06	2	628.53	20.03	.000
Error (Similarity)	1445.61	46	31.426		
Position	19328.65	2.691	7182.834	181.49	.000
Error (Position)	2449.52	61.892	39.577		
Similarity*Position	371.83	10	37.183	5.32	.000
Error (Similarity*Position)	1606.84	230	6.986		

Table B 2: Post-hoc paired samples *t*-tests on the mean differences in performance between the similarity conditions when scored using the correct-in-position measure for rhyming words (Experiment 1).

Source	***t*-value**	***df***	**2-tail Sig**
Dissimilar vs. Rhyming	4.126	23	.000
Dissimilar vs. Similar	6.609	23	.000
Rhyming vs. Similar	1.391	23	.178

Table B 3: One-way repeated measures ANOVA on the mean number of correct responses at an item recall level for rhyming words (Experiment 1).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	5557.19	1.583	3511.52	28.42	.000
Error (Similarity)	4496.81	36.399	123.54		

Table B 4: Post-hoc paired samples *t*-test on the mean differences in performance between the similarity conditions when scored at an item recall level for rhyming words (Experiment 1).

Source	*t*-value	*df*	2-tail Sig
Rhyming vs. Dissimilar	4.885	23	.000
Dissimilar vs. Similar	1.063	23	.299
Rhyming vs. Similar	8.626	23	.000

Table B 5: One-way repeated measures ANOVA on the proportion correct as a function of the number of items recalled using the order accuracy measure for rhyming words (Experiment 1).

Source	SS	*df*	MS	*F*	*P*
Similarity	.784	2	.392	76.023	.000
Error (Similarity)	.237	46	.0052		

Table B 6: Post-hoc paired samples *t*-tests on the proportion correct as a function of the number of items recalled using the order accuracy measure for rhyming words (Experiment 1).

Source	*t*-value	*df*	2-tail Sig
Dissimilar vs. Rhyming	11.330	23	.000
Dissimilar vs. Similar	10.316	23	.000
Similar vs. Rhyming	2.622	23	.015

Table B 7: Repeated measures 3 x 6 ANOVA on the mean number of correct responses using the correct-in-position measure for CV_ lists (Experiment 2).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	2145.505	2	1072.752	34.996	.000
Error (Similarity)	1410.051	46	30.653		
Position	20331.373	3.017	6738.365	238.865	.000
Error (Position)	1957.683	69.397	28.210		
Similarity*Position	293.384	10	29.338	4.192	.000
Error (Similarity*Position)	1609.727	230	6.999		

Table B 8: Post-hoc paired samples *t*-tests on the mean differences in performance between the similarity conditions when scored using the correct-in-position measure for CV_ lists (Experiment 2).

Source	***t*-value**	***df***	**2-tail Sig**
Dissimilar vs. CV_ Lists	7.222	23	.000
Dissimilar vs. Similar	4.876	23	.000
Similar vs. CV_ Lists	4.278	23	.000

Table B 9: One-way repeated measures ANOVA on the mean number of correct responses at an item recall level for CV_ lists (Experiment 2).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	497.194	1.424	349.231	2.517	.111
Error (Similarity)	4542.806	32.745	138.734		

Table B 10: One-way repeated measures ANOVA on the proportion correct as a function of the number of items recalled using the order accuracy measure for CV_ lists (Experiment 2).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	.867	2	.433	83.061	.000
Error (Similarity)	.240	46	.0052		

Table B 11: Post-hoc paired samples *t*-tests on the proportion correct as a function of the number of items recalled using the order accuracy measure for CV_ lists (Experiment 2).

Source	***t*-value**	***df***	**2-tail Sig**
Dissimilar vs. CV_ Lists	12.365	23	.000
Dissimilar vs. Similar	6.591	23	.000
Similar vs. CV_ Lists	6.659	23	.000

Table B 12: Repeated measures 3 x 6 ANOVA on the mean number of correct responses using the correct-in-position measure for C_C lists (Experiment 3).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	2042.347	2	1021.174	32.172	.000
Error (Similarity)	1460.097	46	31.741		
Position	15986.333	2.974	5374.739	183.313	.000
Error (Position)	2005.778	68.410	29.320		
Similarity*Position	320.236	10	32.024	5.716	.000
Error (Similarity*Position)	1288.653	230	5.603		

Table B 13: Post-hoc paired samples *t*-tests on the mean differences in performance between the similarity conditions when scored using the correct-in-position measure for C_C lists (Experiment 3).

Source	*t*-value	*df*	2-tail Sig
Dissimilar vs. C_C Lists	4.778	23	.000
Dissimilar vs. Similar	7.507	23	.000
C_C Lists vs. Similar	3.384	23	.003

Table B 14: One-way repeated measures ANOVA on the mean number of correct responses at an item recall level for C_C lists (Experiment 3).

Source	SS	*df*	MS	*F*	*P*
Similarity	8294.361	1.576	5262.675	37.316	.000
Error (Similarity)	5112.306	36.250	141.030		

Table B 15: Post-hoc paired samples *t*-test on the mean differences in performance between the similarity conditions when scored at an item recall level for C_C lists (Experiment 3).

Source	*t*-value	*df*	2-tail Sig
C_C Lists vs. Dissimilar	2.043	23	.053
Dissimilar vs. Similar	5.757	23	.000
C_C Lists vs. Similar	11.662	23	.000

Table B 16: One-way repeated measures ANOVA on the proportion correct as a function of the number of items recalled using the order accuracy measure for C_C lists (Experiment 3).

Source	SS	*df*	MS	*F*	*P*
Similarity	.534	2	.267	48.578	.000
Error (Similarity)	.253	46	.0055		

Table B 17: Post-hoc paired samples *t*-tests on the proportion correct as a function of the number of items recalled using the order accuracy measure for C_C lists (Experiment 3).

Source	*t*-value	*df*	2-tail Sig
Dissimilar vs. C_C Lists	10.566	23	.000
Dissimilar vs. Similar	7.314	23	.000
Similar vs. C_C Lists	1.646	23	.113

Table B 18: One-way between subjects ANOVA on the proportion of recombination errors across the three experiments for the phonemically similar conditions

Source	SS	*df*	MS	*F*	*P*
Recombinations	.075	2	.038	12.914	.000
Error (Recombinations)	.201	69	.003		

Table B 19: Post-hoc independent samples *t*-tests on the proportion of recombination errors as a function of the phonemically similar condition.

Source	*t*-value	*df*	2-tail Sig
Exp 2 (C_C & _VC) vs. Exp 1 (CV_ & C_C)	2.696	46	.010
Exp 3 (CV_ & _VC) vs. Exp 1 (CV_ & C_C)	4.653	46	.000
Exp 3 (CV_ & _VC) vs Exp 2 (C_C & _VC)	2.672	46	.010

Appendix C – Stimulus Lists and IPA codes used in Chapter 5 (Study Two)

Table C1. *Rhyming Nonword Set Used for Experiment 1*

Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes
pame	peɪm	vame	veɪm	yame	jeɪm	wame	weɪm	zame	zeɪm
hace	heɪs	zace	zeɪs	tace	teɪs	nace	neɪs	wace	weɪs
tait	teɪt	nait	neɪt	vait	veɪt	yait	jeɪt	zait	zeɪt
wayb	weɪb	hayb	heɪb	tayb	teɪb	fayb	feɪb	gayb	geɪb
zile	zaɪl	shile	ʃaɪl	yile	jaɪl	hile	haɪl	chile	tʃaɪl
shied	ʃaɪd	kied	kaɪd	mied	maɪd	zied	zaɪd	yied	yaɪd
zine	zaɪn	yine	jaɪn	hine	haɪn	gine	gaɪn	kine	kaɪn
boke	bəʊk	doke	dəʊk	goke	gəʊk	noke	nəʊk	hoke	həʊk
gome	gəʊm	wome	wəʊm	lome	ləʊm	pome	pəʊm	bome	bəʊm
yone	jəʊn	vone	vəʊn	pone	pəʊn	chone	tʃəʊn	wone	wəʊn
wole	wəʊl	chole	tʃəʊl	nole	nəʊl	lole	ləʊl	zole	zəʊl
pag	pæg	kag	kæg	vag	væg	chag	tʃæg	yag	jæg
kang	kæŋ	mang	mæŋ	nang	næŋ	dang	dæŋ	wang	wæŋ
bam	bæm	nam	næm	fam	fæm	cham	tʃæm	vam	væm
fap	fæp	vap	væp	wap	wæp	bap	bæp	dap	dæp
gin	gɪn	zin	zɪn	hin	hɪn	nin	nɪn	min	mɪn
bip	bɪp	gip	gɪp	mip	mɪp	fip	fɪp	vip	vɪp
hig	hɪg	chig	tʃɪg	nig	nɪg	vig	vɪg	yig	jɪg
ning	nɪŋ	hing	hɪŋ	ching	tʃɪŋ	ming	mɪŋ	fing	fɪŋ
geed	gi:d	yeed	ji:d	meed	mi:d	zeed	mi:d	veed	vi:d
teep	ti:p	veep	vi:p	feep	fi:p	zeep	zi:p	yeep	ji:p
deek	di:k	heek	hi:k	neek	ni:k	feek	fi:k	yeek	ji:k
leeb	li:b	teeb	ti:b	heeb	hi:b	geeb	gi:b	deeb	di:b
mun	mʌn	yun	jʌn	vun	vʌn	zun	zʌn	lun	lʌn
wug	wʌg	zug	zʌg	vug	vʌg	shug	ʃʌg	kug	kʌg
lub	lʌb	fub	fʌb	gub	gʌb	shub	ʃʌb	mub	mʌb
wut	wʌt	vut	vʌt	zut	zʌt	yut	jʌt	chut	tʃʌt
zock	zɒk	yock	jɒk	gock	gɒk	fock	fɒk	vock	vɒk
shog	ʃɒg	pog	pɒg	chog	tʃɒg	yog	jɒg	mog	mɒg
chot	tʃɒt	zot	zɒt	mot	mɒt	fot	fɒt	vot	vɒt

Table C2. *Common CV_ Nonword Set Used for Experiment 2*

Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes
mayp	meɪp	mayb	meɪb	mayg	meɪg	mayf	meɪf	maych	meɪtʃ
says	seɪs	sayp	seɪp	sayg	seɪg	sayng	seɪŋ	saych	seɪtʃ
waysh	weɪʃ	ways	weɪs	waych	weɪtʃ	waym	weɪm	wayb	weɪb
raych	reɪtʃ	rayb	reɪb	rayg	reɪg	rayng	reɪŋ	raym	reɪm
lieg	laɪg	lieng	laɪŋ	liech	laɪtʃ	liesh	laɪʃ	liedge	laɪdʒ
riech	raɪtʃ	riedge	raɪdʒ	riesh	raɪʃ	riev	raɪv	rieng	raɪŋ
tiech	taɪtʃ	ties	taɪs	tieng	taɪŋ	tieg	taɪg	tiev	taɪv
wieb	waɪb	wiesh	waɪʃ	wieng	waɪŋ	wies	waɪs	wiem	waɪm
birsh	bɜ:ʃ	birf	bɜ:f	birng	bɜ:ŋ	birm	bɜ:m	birb	bɜ:b
pirsh	pɜ:ʃ	pirg	pɜ:g	pirng	pɜ:ŋ	pirb	pɜ:b	pirf	pɜ:f
kav	kæv	kas	kæs	kaz	kæz	kang	kæŋ	kag	kæg
hab	hæb	han	hæn	haf	hæf	hadge	hædʒ	has	hæs
mab	mæb	maf	mæf	mav	mæv	maz	mæz	madge	mædʒ
wab	wæb	wan	wæn	wadge	wædʒ	wach	wætʃ	wav	wæv
lig	lɪg	lish	lɪʃ	lidge	lɪdʒ	lif	lɪf	lis	lɪs
pish	pɪʃ	piv	pɪv	pim	pɪm	pidge	pɪdʒ	pib	pɪb
bish	bɪʃ	bim	bɪm	biv	bɪv	bis	bɪs	bidge	bɪdʒ
seeb	sɪ:b	seech	sɪ:tʃ	seesh	sɪ:ʃ	seef	sɪ:f	seeg	sɪ:g
weesh	wɪ:ʃ	weech	wɪ:tʃ	weef	wɪ:f	weem	wɪ:m	weeg	wɪ:g
beev	bɪ:v	beesh	bɪ:ʃ	bees	bɪ:s	beeg	bɪ:g	beeb	bɪ:b
kudge	kʌdʒ	kung	kʌŋ	kuv	kʌv	kuch	kʌtʃ	kug	kʌg
hus	hʌs	huz	hʌz	hudge	hʌdʒ	hup	hʌp	huv	hʌv
mup	mʌp	muv	mʌv	mus	mʌs	muz	mʌz	mudge	mʌdʒ
ruch	rʌtʃ	rudge	rʌdʒ	rus	rʌs	ruv	rʌv	rup	rʌp
losh	lɒʃ	loch	lɒtʃ	lon	lɒn	lom	lɒm	lof	lɒf
kodge	lɒdʒ	kosh	lɒʃ	kom	lɒm	kov	lɒv	koch	lɒtʃ
sog	sɒg	som	sɒm	sodge	sɒdʒ	son	sɒn	sosh	sɒʃ
porv	pɔ:v	porb	pɔ:b	porg	pɔ:g	porf	pɔ:f	pors	pɔ:s
worb	wɔ:b	worg	wɔ:g	worch	wɔ:tʃ	worv	wɔ:v	wors	wɔ:s
harb	hɑ:b	harn	hɑ:n	hardge	hɑ:dʒ	hars	hɑ:s	harz	hɑ:z

Table C3. *Common C_C Nonword Set Used for Experiment 3*

Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes
bech	bɛtʃ	barch	bɑːtʃ	borch	bɔːtʃ	baych	beɪtʃ	biech	baɪtʃ
bidge	bɪdʒ	bedge	bɛdʒ	beedge	biːdʒ	baydge	beɪdʒ	biedge	baɪdʒ
bish	bɪʃ	besh	bvʃ	biesh	baɪʃ	barsh	bɑːʃ	baysh	beɪʃ
biv	bɪv	beev	biːv	barv	bɑːv	borv	bɔːv	bayv	beɪv
didge	dɪdʒ	dedge	dɛdʒ	dadge	dædʒ	deedge	diːdʒ	dardge	dɑːdʒ
div	dɪv	dav	dæv	dov	dɒv	darv	dɑːv	deev	diːv
diz	dɪz	dez	dɛz	daz	dæz	doz	dɒz	darz	dɑːz
has	hæs	hes	hɛs	hees	hiːs	hays	heɪs	hies	haɪs
haf	hæf	heef	hiːf	hirf	hɜːf	hayf	heɪf	hief	haɪf
han	hæn	hon	hɒn	hirn	hɜːn	hayn	heɪn	hien	haɪn
hom	hɒm	heem	hiːm	hirm	hɜːm	haym	heɪm	hiem	haɪm
chaf	tʃæf	chof	tʃɒf	charf	tʃɑːf	chirf	tʃɜːf	chief	tʃaɪf
yeb	jɛb	yeeb	jiːb	yob	jɒb	yarb	jɑːb	yieb	jaɪb
yeg	jɛg	yeeg	jiːg	yog	jɒg	yarg	jɑːg	yieg	jaɪg
cheg	tʃɛg	chog	tʃɒg	charg	tʃɑːg	chirg	tʃɜːg	chieg	tʃaɪg
yek	jɛk	yeek	jiːk	yock	jɒk	yark	jɑːk	yiek	jaɪk
yem	jɛm	yeem	jiːm	yom	jɒm	yarm	jɑːm	yiem	jaɪm
larv	lɑːv	lev	lɛv	lorv	lɔːv	lav	læv	layv	leɪv
lef	lɛf	laf	læf	lorf	lɔːf	lirf	lɜːf	layf	leɪf
lidge	lɪdʒ	liedge	laɪdʒ	lordge	lɔːdʒ	ladge	lædʒ	laydge	leɪdʒ
lish	lɪʃ	lesh	lɛʃ	liesh	laɪʃ	lirsh	lɜːʃ	laysh	leɪʃ
res	rɛs	ras	ræs	rars	rɑːs	rors	rɔːs	rirs	rɜːs
rin	rɪn	reen	riːn	rarn	rɑːn	rorn	rɔːn	rirn	rɜːn
riz	rɪz	rez	rɛz	reez	riːz	rarz	rɑːz	rirz	rɜːz
tas	tæs	tus	tʌs	tors	tɔːs	tays	teɪs	ties	taɪs
teg	tɛg	targ	tɑːg	torg	tɔːg	tayg	teɪg	tieg	taɪg
tem	tɛm	tam	tæm	tum	tʌm	tarm	tɑːm	torm	tɔːm
wadge	wædʒ	wudge	wʌdʒ	weedge	wiːdʒ	wardge	wɑːdʒ	wiedge	waɪdʒ
wem	wɛm	wum	wʌm	weem	wiːm	warm	wɑːm	wiem	waɪm
wesh	wɛʃ	wash	wæʃ	wush	wʌʃ	waysh	weɪʃ	wiesh	waɪʃ

Appendix D – ANOVA Tables for the major statistical analyses performed for Chapter 5 (Study Two)

Table D 1: One-way repeated measures ANOVA on the mean number of correct responses using the correct-in-position measure for rhyming nonwords (Experiment 1).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	3251.583	2	1625.792	16.672	.000
Error (Similarity)	4485.750	46	97.516		

Table D 2: Post-hoc paired samples *t*-tests on the mean differences in performance between the similarity conditions when scored using the correct-in-position measure for rhyming nonwords (Experiment 1).

Source	***t*-value**	***df***	**2-tail Sig**
Rhyming vs. Dissimilar	5.041	23	.000
Similar vs. Dissimilar	4.591	23	.000
Rhyming vs. Similar	1.250	23	.224

Table D 3: One-way repeated measures ANOVA on the mean number of correct responses at an item recall level for rhyming nonwords (Experiment 1).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	21139.083	2	10569.542	118.636	.000
Error (Similarity)	4098.250	46	89.092		

Table D 4: Post-hoc paired samples *t*-test on the mean differences in performance between the similarity conditions when scored at an item recall level for rhyming nonwords (Experiment 1).

Source	*t*-value	*df*	2-tail Sig
Rhyming vs. Dissimilar	13.794	23	.000
Similar vs. Dissimilar	9.055	23	.000
Rhyming vs. Similar	7.123	23	.000

Table D 5: One-way repeated measures ANOVA on the proportion correct as a function of the number of items recalled using the order accuracy measure for rhyming nonwords (Experiment 1).

Source	SS	*df*	MS	*F*	*P*
Similarity	.714	2	.357	85.938	.000
Error (Similarity)	.191	46	.0042		

Table D 6: Post-hoc paired samples *t*-tests on the proportion correct as a function of the number of items recalled using the order accuracy measure for rhyming nonwords (Experiment 1).

Source	*t*-value	*df*	2-tail Sig
Dissimilar vs. Rhyming	12.454	23	.000
Dissimilar vs. Similar	8.439	23	.000
Similar vs. Rhyming	5.334	23	.000

Table D 7: One-way repeated measures ANOVA on the mean number of correct responses using the correct-in-position measure for CV_ nonword lists (Experiment 2).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	5054.694	2	2527.347	40.603	.000
Error (Similarity)	2863.306	46	62.246		

Table D 8: Post-hoc paired samples *t*-tests on the mean differences in performance between the similarity conditions when scored using the correct-in-position measure for CV_ lists of nonwords (Experiment 2).

Source	***t*-value**	***df***	**2-tail Sig**
Dissimilar vs. CV_ Lists	.752	23	.460
Similar vs. Dissimilar	6.456	23	.000
Similar vs. CV_ Lists	8.368	23	.000

Table D 9: One-way repeated measures ANOVA on the mean number of correct responses at an item recall level for CV_ lists of nonwords (Experiment 2).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	10755.583	2	5377.792	83.751	.000
Error (Similarity)	2983.750	46	64.212		

Table D 10: Post-hoc paired samples *t*-test on the mean differences in performance between the similarity conditions when scored at an item recall level for CV_ lists of nonwords (Experiment 2).

Source	t-value	*df*	2-tail Sig
CV_ Lists vs. Dissimilar	9.788	23	.000
Similar vs. Dissimilar	11.295	23	.000
Similar vs. CV_ Lists	2.744	23	.012

Table D 11: One-way repeated measures ANOVA on the proportion correct as a function of the number of items recalled using the order accuracy measure for CV_ lists of nonwords (Experiment 2).

Source	SS	*df*	MS	*F*	*P*
Similarity	1.241	2	.621	95.719	.000
Error (Similarity)	.298	46	.0065		

Table D 12: Post-hoc paired samples *t*-tests on the proportion correct as a function of the number of items recalled using the order accuracy measure for CV_ lists of nonwords (Experiment 2).

Source	*t*-value	*df*	2-tail Sig
Dissimilar vs. CV_ Lists	13.142	23	.000
Dissimilar vs. Similar	5.358	23	.000
Similar vs. CV_ Lists	8.833	23	.000

Table D 13: One-way repeated measures ANOVA on the mean number of correct responses using the correct-in-position measure for C_C lists of nonwords (Experiment 3).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	20209.000	2	10104.500	74.270	.000
Error (Similarity)	6258.333	46	136.051		

Table D 14: Post-hoc paired samples *t*-tests on the mean differences in performance between the similarity conditions when scored using the correct-in-position measure for C_C lists of nonwords (Experiment 3).

Source	***t*-value**	***df***	**2-tail Sig**
C_C Lists Vs. Dissimilar	11.193	23	.000
Similar Vs. Dissimilar	2.372	23	.026
C_C Lists Vs. Similar	8.157	23	.000

Table D 15: One-way repeated measures ANOVA on the mean number of correct responses at an item recall level for C_C lists of nonwords (Experiment 3).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	35423.444	2	17711.722	179.991	.000
Error (Similarity)	4526.556	46	98.403		

Table D 16: Post-hoc paired samples *t*-test on the mean differences in performance between the similarity conditions when scored at an item recall level for C_C lists of nonwords (Experiment 3).

Source	*t*-value	*df*	2-tail Sig
C_C Lists vs. Dissimilar	18.159	23	.000
Similar vs. Dissimilar	8.275	23	.000
C_C Lists vs. Similar	11.084	23	.000

Table D 17: One-way repeated measures ANOVA on the proportion correct as a function of the number of items recalled using the order accuracy measure for C_C lists of nonwords (Experiment 3).

Source	SS	*df*	MS	*F*	*P*
Similarity	.741	1.553	.477	55.875	.000
Error (Similarity)	.305	35.729	.0085		

Table D 18: Post-hoc paired samples *t*-tests on the proportion correct as a function of the number of items recalled using the order accuracy measure for C_C lists of nonwords (Experiment 3).

Source	*t*-value	*df*	2-tail Sig
Dissimilar vs. C_C Lists	8.381	23	.000
Dissimilar vs. Similar	10.944	23	.000
C_C Lists vs. Similar	3.343	23	.003

Table D 19: One-way between subjects ANOVA across the three experiments on the scores obtained using the item recall measure of performance for the three phonemically dissimilar conditions.

Source	**SS**	***df***	**MS**	***F***	***P***
Dissimilar	1387.750	2	693.875	3.458	.037
Error (Similarity)	13846.25	69	200.670		

Table D 20: Post-hoc independent samples *t*-tests on the number of items recall using the item recall measure of performance for the phonemically dissimilar lists of nonwords.

Source	***t*-value**	***df***	**2-tail Sig**
Experiment 1 vs. Experiment 2	1.310	46	.197
Experiment 1 vs. Experiment 3	2.472	46	.017
Experiment 2 vs. Experiment 3	1.408	46	.166

Table D 21: One-way between subjects ANOVA across the three experiments on the scores obtained using the order accuracy measure of performance for the three phonemically dissimilar conditions.

Source	**SS**	***df***	**MS**	***F***	***P***
Dissimilar	.007	2	.003	.905	.409
Error (Similarity)	.265	69	.004		

Table D 22: One-way between subjects ANOVA across the three experiments on the scores obtained using the order accuracy measure of performance for the three phonemically similar conditions (i.e., _VC, CV_ and C_C lists).

Source	**SS**	***df***	**MS**	***F***	***P***
Similarity	.436	2	.218	19.244	.000
Error (Similarity)	.782	69	.011		

Table D 23: Post-hoc independent samples *t*-tests on the scores obtained using the order accuracy measure of performance for the phonemically similar conditions (i.e., VC, CV and C_C lists).

Source	***t*-value**	***df***	**2-tail Sig**
_VC vs. CV_	2.395	46	.021
_VC vs. C_C	-3.854	46	.000
CV_ vs. C_C	-6.064	46	.000

Appendix E – Stimulus lists and IPA codes used for the Nonword Experiments in Chapter 6 (Study Three)

Table E1. *Rhyming Nonword Set Used for Experiment 1*

Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes
vame	veɪm	hame	heɪm	yame	jeɪm	wame	weɪm	zame	zeɪm	rame	reɪm
hace	heɪs	zace	zeɪs	tace	teɪs	nace	neɪs	wace	weɪs	yace	jeɪs
taych	peɪtʃ	naych	neɪtʃ	waych	weɪtʃ	yaych	jeɪtʃ	gaych	geɪtʃ	paych	peɪtʃ
payb	peɪb	fayb	feɪb	nayb	neɪb	wayb	weɪb	hayb	heɪb	rayb	reɪb
shiep	ʃaɪp	fiep	faɪp	kiep	kaɪp	diep	daɪp	liep	laɪp	miep	maɪp
shied	ʃaɪd	kied	kaɪd	mied	maɪd	zied	zaɪd	yied	jaɪd	chied	tʃaɪd
chiet	tʃaɪt	yiet	jaɪt	giet	gaɪt	ziet	zaɪt	viet	vaɪt	diet	daɪt
fieb	faɪb	hieb	haɪb	mieb	maɪb	gieb	gaɪb	kieb	kaɪb	lieb	laɪb
boke	bəʊk	doke	dəʊk	shoke	ʃəʊk	noke	nəʊk	hoke	həʊk	roke	rəʊk
gome	gəʊm	pome	pəʊm	lome	ləʊm	chome	tʃəʊm	bome	bəʊm	shome	ʃəʊm
yone	jəʊn	vone	vəʊn	pone	pəʊn	chone	tʃəʊn	wone	wəʊn	rone	rəʊn
wole	wəʊl	chole	tʃəʊl	nole	nəʊl	lole	ləʊl	vole	vəʊl	yole	jəʊl
gin	gɪn	zin	zɪn	hin	hɪn	rin	rɪn	min	mɪn	vin	vɪn
shid	ʃɪd	gid	gɪd	nid	nɪd	fid	fɪd	vid	vɪd	zid	zɪd
hig	hɪg	chig	tʃɪg	nig	nɪg	vig	vɪg	kig	kɪg	shig	ʃɪg
ning	nɪŋ	hing	hɪŋ	ging	gɪŋ	ming	mɪŋ	fing	fɪŋ	shing	ʃɪŋ
neech	ni:tʃ	geech	gi:tʃ	heech	hi:tʃ	veech	vi:tʃ	cheech	tʃi:tʃ	yeech	ji:tʃ
teep	ti:p	veep	vi:p	feep	fi:p	geep	gi:p	yeep	ji:p	meep	mi:p
deek	di:k	heek	hi:k	neek	ni:k	feek	fi:k	yeek	ji:k	veek	vi:k
weeb	wi:b	teeb	ti:b	heeb	hi:b	cheeb	tʃi:b	veeb	vi:b	reeb	ri:b
boz	bɒz	moz	mɒz	toz	tɒz	voz	vɒz	zoz	zɒz	poz	pɒz
shog	ʃɒg	pog	pɒg	chog	tʃɒg	tog	tɒg	mog	mɒg	zog	zɒg
chot	tʃɒt	zot	zɒt	mot	mɒt	fot	fɒt	vot	vɒt	bot	bɒt
mun	mʌn	yun	jʌn	kun	kʌn	zun	zʌn	lun	lʌn	hun	hʌn
wug	wʌg	zug	zʌg	vug	vʌg	shug	ʃʌg	kug	kʌg	gug	gʌg
lub	lʌb	fub	fʌb	gub	gʌb	shub	ʃʌb	mub	mʌb	wub	wʌb
wut	wʌt	vut	vʌt	zut	zʌt	yut	jʌt	chut	tʃʌt	fut	fʌt
vang	væŋ	mang	mæŋ	nang	næŋ	dang	dæŋ	wang	wæŋ	shang	ʃæŋ
bam	bæm	nam	næm	fam	fæm	cham	tʃæm	vam	væm	gam	gæm
fap	fæp	vap	væp	wap	wæp	bap	bæp	dap	dæp	shap	ʃæp

Table E2*. Common CV_ Nonword Set Used for Experiment 2*

Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes
mayp	meɪp	mayv	meɪv	mayg	meɪg	mayf	meɪf	maych	meɪtʃ	maysh	meɪʃ
says	seɪs	sayp	seɪp	sayg	seɪg	sayng	seɪŋ	saych	seɪtʃ	saysh	seɪʃ
waysh	weɪʃ	ways	weɪs	waych	weɪtʃ	waym	weɪm	wayb	weɪb	wayg	weɪg
raych	reɪtʃ	raysh	reɪʃ	rayg	reɪg	rayng	reɪŋ	raym	reɪm	rayf	reɪf
lieg	laɪg	lieng	laɪŋ	liech	laɪtʃ	liesh	laɪʃ	liedge	laɪdʒ	lieb	laɪb
riech	raɪtʃ	riedge	raɪdʒ	riesh	raɪʃ	riev	raɪv	rieng	raɪŋ	rieg	raɪg
tiech	taɪtʃ	tiesh	taɪʃ	tieng	taɪŋ	tieg	taɪg	tiev	taɪv	tiedge	taɪdʒ
wieb	waɪb	wiesh	waɪʃ	wieng	waɪŋ	wies	waɪs	wiem	waɪm	wieg	waɪg
birsh	bɜ:ʃ	birf	bɜ:f	birng	bɜ:ŋ	birm	bɜ:m	birv	bɜ:v	birdge	bɜ:dʒ
pirsh	pɜ:ʃ	pirg	pɜ:g	pirng	pɜ:ŋ	pirb	pɜ:b	pirf	pɜ:f	pirp	pɜ:p
kav	kæv	kas	kæs	kaz	kæz	kang	kæŋ	kag	kæg	kadge	kædʒ
taf	tæf	tadge	tædʒ	tas	tæs	tach	tætʃ	tam	tæm	taz	tæz
nang	næŋ	nav	næv	naz	næz	nadge	nædʒ	nam	næm	nas	næs
waz	wæz	waf	wæf	wadge	wædʒ	wach	wætʃ	wav	wæv	was	wæs
lig	lɪg	lish	lɪʃ	lidge	lɪdʒ	lif	lɪf	lis	lɪs	lib	lɪb
pish	pɪʃ	piv	pɪv	pim	pɪm	pidge	pɪdʒ	pib	pɪb	pif	pɪf
bish	bɪʃ	bim	bɪm	biv	bɪv	bis	bɪs	bidge	bɪdʒ	bing	bɪŋ
seeb	sɪ:b	seech	sɪ:tʃ	seesh	sɪ:ʃ	seef	sɪ:f	seeg	sɪ:g	seev	sɪ:v
weesh	wɪ:ʃ	wees	wɪ:s	weef	wɪ:f	weem	wɪ:m	weeg	wɪ:g	weedge	wɪ:dʒ
beev	bɪ:v	beesh	bɪ:ʃ	bees	bɪ:s	beeg	bɪ:g	beeb	bɪ:b	beedge	bɪ:dʒ
kuz	kʌz	kung	kʌŋ	kuv	kʌv	kuch	kʌtʃ	kug	kʌg	kun	kʌn
lus	lʌs	lum	lʌm	luz	lʌz	lub	lʌb	lun	lʌn	luch	lʌtʃ
mup	mʌp	muv	mʌv	mus	mʌs	mun	mʌn	mudge	mʌdʒ	muz	mʌz
ruch	rʌtʃ	rudge	rʌdʒ	rus	rʌs	ruv	rʌv	rup	rʌp	ruz	rʌz
loz	lɒz	loch	lɒtʃ	lon	lɒn	lom	lɒm	lof	lɒf	lov	lɒv
hodge	hɒdʒ	hon	hɒn	hob	hɒb	hoch	hɒtʃ	hos	hɒs	hoz	hɒz
tog	tɒg	tov	tɒv	todge	tɒdʒ	tosh	tɒʃ	toch	tɒtʃ	toz	tɒz
porv	pɔ:v	porp	pɔ:p	porg	pɔ:g	porf	pɔ:f	pors	pɔ:s	porm	pɔ:m
worb	wɔ:b	worg	wɔ:g	worch	wɔ:tʃ	worv	wɔ:v	wors	wɔ:s	worsh	wɔ:ʃ
harb	hɑ:b	harn	hɑ:n	hardge	hɑ:dʒ	hars	hɑ:s	harz	hɑ:z	harch	hɑ:tʃ

Table E3. *Common C_C Nonword Set Used for Experiment 3*

Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes	Stimuli	IPA Codes
bech	bɛtʃ	barch	bɑ:tʃ	borch	bɔ:tʃ	baych	beɪtʃ	biech	baɪtʃ	buch	bʌʃ
bidge	bɪdʒ	bedge	bɛdʒ	birdge	bɜ:dʒ	baydge	beɪdʒ	biedge	baɪdʒ	bodge	bɒdʒ
bish	bɪʃ	borsh	bɔ:ʃ	biesh	baɪʃ	barsh	bɑ:ʃ	baysh	beɪʃ	bush	bʌʃ
biv	bɪv	bav	bæv	barv	bɑ:v	birv	bɜ:v	bayv	beɪv	bov	bɒv
chaf	tʃæf	chof	tʃɒf	charf	tʃɑ:f	chirf	tʃɜ:f	chief	tʃaɪf	chorf	tʃɔ:f
cheg	tʃɛg	chog	tʃɒg	charg	tʃɑ:g	chirg	tʃɜ:g	chag	tʃæg	chorg	tʃɔ:g
didge	dɪdʒ	dedge	dɛdʒ	dadge	dædʒ	dudge	dʌdʒ	dardge	dɑ:dʒ	diedge	daɪdʒ
div	dɪv	dav	dæv	dov	dɒv	darv	dɑ:v	dirv	dɜ:v	dorv	dɔ:v
diz	dɪz	dez	dɛz	daz	dæz	doz	dɒz	darz	dɑ:z	dirz	dɜ:z
gim	gɪm	girm	gɜ:m	garm	gɑ:m	gorm	gɔ:m	geem	gi:m	gem	gɛm
gin	gɪn	geen	gi:n	garn	gɑ:n	gen	gɛn	girn	gɜ:n	gon	gɒn
haf	hæf	heef	hi:f	hirf	hɜ:f	hayf	heɪf	hief	haɪf	hof	hɒf
han	hæn	hon	hɒn	hirn	hɜ:n	hayn	heɪn	hien	haɪn	harn	hɑ:n
hees	hi:s	hos	hɒs	hars	hɑ:s	hays	heɪs	hies	haɪs	hus	hʌs
hom	hɒm	heem	hi:m	hirm	hɜ:m	haym	heɪm	hiem	haɪm	horm	hɔ:m
larv	lɑ:v	lev	lɛv	lorv	lɔ:v	lav	læv	layv	leɪv	lov	lɒv
lef	lɛf	laf	læf	lof	lɒf	lirf	lɜ:f	layf	leɪf	luf	lʌf
lidge	lɪdʒ	liedge	laɪdʒ	lordge	lɔ:dʒ	ladge	lædʒ	laydge	leɪdʒ	ludge	lʌdʒ
lish	lɪʃ	lesh	lɛʃ	liesh	laɪʃ	lirsh	lɜ:ʃ	laysh	leɪʃ	larsh	lɑ:ʃ
res	rɛs	ras	ræs	rars	rɑ:s	rors	rɔ:s	rirs	rɜ:s	rus	rʌs
riz	rɪz	rez	rɛz	raz	ræz	rarz	rɑ:z	rirz	rɜ:z	ruz	rʌz
tas	tæs	tus	tʌs	tors	tɔ:s	tays	teɪs	ties	taɪs	tees	ti:s
tog	tɒg	targ	tɑ:g	torg	tɔ:g	tayg	teɪg	tieg	taɪg	teeg	ti:g
wadge	wædʒ	wudge	wʌdʒ	wirdge	wɜ:dʒ	wardge	wɑ:dʒ	wiedge	waɪdʒ	wodge	wɒdʒ
wem	wɛm	wum	wʌm	weem	wi:m	warm	wɑ:m	wiem	waɪm	waym	weɪm
wesh	wɛʃ	wash	wæʃ	wush	wʌʃ	waysh	weɪʃ	wiesh	waɪʃ	warsh	wɑ:ʃ
yeb	jɛb	yeeb	ji:b	yarb	jɑ:b	yab	jæb	yieb	jaɪb	yirb	jɜ:b
yeg	jɛg	yeeg	ji:g	yog	jɒg	yarg	jɑ:g	yag	jæg	yayg	jeɪg
yek	jɛk	yeek	ji:k	yock	jɒk	yark	jɑ:k	yiek	jaɪk	yayk	jeɪk
yem	jɛm	yeem	ji:m	yom	jɒm	yarm	jɑ:m	yim	jɪm	yirm	jɜ:m

Appendix F – ANOVA Tables for the major statistical analyses performed for Chapter 6 (Study Three)

Table F 1: Mixed design 2 (lexicality) x 3 (similarity) ANOVA on the mean d-prime (*d'*) value for recognition judgments for Experiment 1.

Source	**SS**	***df***	**MS**	***F***	***P***
Lexicality (Between)	.108	1	.108	.170	.681
Error	44.336	70	.633		
Similarity (Within)	10.160	2	5.080	14.871	.000
Similarity x Lexicality	.114	2	.0057	.164	.849
Error (Similarity)	48.810	140	.349		

Table F 2: Post-hoc paired samples *t*-tests on the mean *d'* value for recognition judgements when the stimuli were phonemically dissimilar, similar or rhymed (Experiment 1).

Source	***t*-value**	***df***	**2-tail Sig**
Dissimilar vs. Rhyme	4.602	71	.000
Dissimilar vs. Similar	4.246	71	.000
Similar vs. Rhyming	.843	71	.402

Table F 3: Mixed design 2 (lexicality) x 3 (similarity) ANOVA on the mean d-prime (*d'*) value for recognition judgments for Experiment 2.

Source	**SS**	***df***	**MS**	***F***	***P***
Lexicality (Between)	.00085	1	.00085	.001	.973
Error	50.892	70	.727		
Similarity (Within)	21.686	2	10.843	30.317	.000
Similarity x Lexicality	.544	2	.272	.761	.469
Error (Similarity)	50.071	140	.358		

Table F 4: Post-hoc paired samples *t*-tests on the mean *d'* value when the stimuli were phonemically dissimilar, similar or shared the CV_ component (Experiment 2).

Source	***t*-value**	***df***	**2-tail Sig**
Dissimilar vs. CV_ Lists	6.873	71	.000
Dissimilar vs. Similar	3.839	71	.000
Similar vs. CV_ Lists	4.385	71	.000

Table F 5: Mixed design 2 (lexicality) x 3 (similarity) ANOVA on the mean d-prime (*d'*) value for recognition judgments for Experiment 3.

Source	**SS**	***df***	**MS**	***F***	***P***
Lexicality (Between)	3.633	1	3.633	5.005	.028
Error	50.812	70	.726		
Similarity (Within)	18.249	2	9.124	28.434	.000
Similarity x Lexicality	1.184	2	.592	1.844	.162
Error (Similarity)	44.927	140	.321		

Table F 6: Post-hoc paired samples *t*-tests on the mean *d'* value when the stimuli were phonemically dissimilar, similar or shared the C_C component (Experiment 3).

Source	***t*-value**	***df***	**2-tail Sig**
Dissimilar vs. C_C Lists	5.176	71	.000
Dissimilar vs. Similar	6.738	71	.000
C_C Lists vs. Similar	2.458	71	.016