Category Properties And The Category-Order Effect

Jordan Schoenherr (psychophysics.lab@gmail.com)
Department of Psychology, Carleton University
1125 Colonel By Drive, Ottawa, ON K1S5B6 Canada

Robert Thomson (rthomson@connect.carleton.ca)
Department of Cognitive Science, Carleton University
1125 Colonel By Drive, Ottawa, ON K1S5B6 Canada

Abstract

Previous research has demonstrated that recall performance is facilitated by the order of presentation of salient stimuli, referred to as the category-order effect (Brooks & Watkins, 1990; Greene & Lasek, 1994). Specifically, recall is higher on the full list when numbers precede words. To extend and clarify this research, Experiment 1 modified Brooks and Watkins' paradigm by presenting items as a unitary string (as opposed to a sequential list) containing four random non-repeating numbers and four letters representing four conditions: random letters, rhyming letters, four letter pseudo-words (consonant-vowel-consonant-vowel strings), and four letter words. In contrast to past results, this experiment found better list recall when words and pseudo-words preceded numbers. Experiment 2 more closely represented Brooks and Watkins' (1990) paradigm by sequentially presenting items in 2 blocks: a block of four numbers and a block of four letters.

Keywords: category-order effect; categorization; list recall; memory; lexicality; working memory

Introduction

Two of the most fundamental aspects to memory are the items stored within it and the order in which these items are encoded and retrieved (Healy, 1974; Lashley, 1951). Although there has been considerable debate over the degree to which this information is inseparable or whether they are in fact independent (Murdock & Von Saal, 1967), there is definitive evidence that the ability to group information together facilitates recall (Ryan, 1969).

The present study examines the conjoint influence of item and order information in terms of the category-order effect (Brooks & Watkins, 1990).

The Category-Order Effect

The category-order effect refers to improved performance on list recall when a set of items from a small, homogeneous category are presented prior to those from a comparatively larger, heterogeneous category (Greene & Lasek, 1994). Category-order effects have been established across several word categories, including when high-frequency words precede average-frequency words (Watkins & Watkins, 1977), semantically-related words precede unrelated words, rhyming words precede non-rhyming words (Brooks & Watkins, 1990), and when numeric digits precede words (Greene & Lasek, 1994; Brooks & Watkins, 1990). However, the category-order effect can be eliminated by decreasing the presentation rate of the stimuli or by using articulatory suppression (Greene & Lasek, 1994).

Greene and Lasek (1994) explored whether the order of recall of the stimuli (i.e. input or output position) was a factor in the category-order effect. To test their hypothesis, participants were asked to either recall items in forward or backward order. In the backward condition, items were to be recalled in reverse order that they were presented in. An additional backward recall condition was also conducted that had participants recall the categories in reverse order, but the items within them in a forward order. Although recall order did affect recall accuracy and overall memory span, both forward and backward recall exhibited a significant category-order effect, with improved recall when more readily categorized items were presented in the first half of the list. Thus, they concluded that input position was the primary determinant of the category-order effect.

Enhancing Memory Capacity

The improved performance on list recall evidenced in the category-order effect is consistent with several results from the broader field of categorization. Landauer and Freeman (1968) demonstrated that the size of a category is inversely related to recognition time. Within categories, common features facilitate recognition of similar members (Feldman, 2003). In terms of lexicality, familiar (i.e. salient) words are more accurately recalled from long-term memory (Waugh & Norman, 1965). Lastly, it has also been shown that digit span is greater than word span (Miller, 1956).

A notable contrasting result is the performance gain exhibited by rhyming words in Brooks and Watkins (1990). Elements that rhyme are generally not as well remembered when presented in earlier portions of a to-be-remembered list, whereas those that are subsequently presented have been known to facilitate list recall (Baddeley, 1986).

Additionally, although the category-order effect seems a natural extension of what is already understood about working memory processes, there have been numerous issues in replicating Watkins' (1977) original findings. As Brooks and Watkins (1990) note, the failure of Klauser and Puckett (1979) is particularly troublesome given that, using Watkins' (1977) stimuli and methodology, no category-order effect was evident in either university students or elderly participants.

Serial Position and the Word-Length Effect

A limitation of category-order studies is the neglect of serial position analyses, especially given that serial position effects are directly concerned with the recall order of subspan lists and have been robustly documented in the memory literature. Two principle findings have been the facilitation for recall of items occupying initial and terminal positions of a list: i.e., primacy and recency effects. These position effects are further altered by the order in which items are recalled. Forward recall is characterized by large primacy and small recency effects, while the reverse trend is observable for backward recall with larger recency effects compared to primacy effects (Atkinson & Shiffrin, 1968). These findings have also been replicated within probe-recognition paradigms (Waugh & Norman, 1965).

Another potential confound within this literature is a lack of control for word length. The word-length effect is the phenomenon where words of relatively longer length have reduced recall rates when compared with words of shorter length (Baddeley, 1986). The word stimuli used in Brooks and Watkins (1990) are related across semantic categories, but vary in their length.

Although it is difficult to imagine creating lists constituted of words of equal length and strongly related, it represents a potential concern given that both Brooks and Watkins' (1990) and Greene and Lasek's (1994) analyses are performed on aggregate scores and do not examine serial position or control for word-length effects.

Lexicality and the Category-Order Effect

A fundamental and transparent finding is that, in memory tasks, words are extremely salient stimuli. Reichler (1969) first demonstrated this in terms of a word superiority effect, where participants were more accurately in identifying a critical letter in the recall phase when that letter was embedded in a word rather than a non-word. This effect potentially arises from a greater number of common features being identified from long-term memory with word stimuli than with non-word stimuli. Alternatively, it may be that it is perceptually simpler to interpret information within word stimuli.

One method of differentiating these hypotheses is through the examination of pseudo-words: stimuli with word-like phonotactic properties but are not part of a subject's lexicon. These words are generally constructed in a consonant-vowel-consonant sequence. Pseudo-words enjoy similar articulation rates to words but have reduced recall when compared with words (Hulme, Maughan, & Brown, 1991).

Present Research

The present study examines several category properties to better understand the mechanisms behind Brooks and Watkins' (1990) and Greene and Lasek's (1994) results that numeric digits preceding words facilitates full-list recall. A corollary of this result is that, at least under their task demands, numeric digits are more salient than words. Experiment 1 modifies Brooks and Watkins' (1990) methodology by dividing stimuli into four letter categories: words, pseudo-words, rhyming letters, and random letters. These stimuli were paired with randomly organized numeric stimuli and were presented as a unit. Word length was controlled and results were scored by position to determine serial position effects as a possible explanadum. Experiment 2 more closely replicated Brooks and Watkins' (1990) original methodology by presenting each stimulus category in a unitary fashion.

Experiment 1

The authors reasoned that a letter sequence representing a word or pseudo-word would be more readily categorized than a sequence of non-repeating numbers, especially if those numbers had no extrinsic meaning such as a specific date or time. Numeric stimuli will still have some salience when presented as a random sequence, as we observe them on a daily basis in the form of monetary values, dates, and other metrics. Under these same conditions, random letters should not exhibit comparable facilitation provided they are not organized into meaningful combinations such as acronyms.

If the category-order effect is simply a function of items sharing membership in the same homogenous category, then numbers and rhyming letters should display significant category-order effects given that they represent a smaller subset of potential digit combinations. Alternatively, if the basis of the category-order effect is indicative of higher-order grouping, such as in lexical categories, category-order effects should arise solely where lexical properties are exhibited. This would result in both word and pseudo-word conditions exhibiting relatively higher recall, with words having the most robust results. These results will not necessarily suggest that the category-order effect is limited to lexical and phonotactic properties, simply that it is indicative of a case where a hierarchy of categories improves chunking strategies.

Method

Participants

Twenty-two undergraduates (15 females, 7 males) participated, and were awarded 1% toward their final grade in an introductory psychology class.

Materials

Each stimulus consisted of eight items divided into two homogenous halves belonging to either the number or letter categories (e.g. BONE1234). Numbers between 1 and 9 were used in a string of 4 non-repeating digits. Numbers expressing historical dates were excluded.

The same method was used to generate the random letter and rhyming letter conditions. A subset of rhyming letters was gathered from the *Handbook of the International Phonetic Association* (1999). Again, any combinations that resembled words or pseudo-words were excluded.

For pseudo-words, consonant-vowel-consonant strings (CVCs) were gathered from Hilgard (1951) and a random non-rhyming vowel was appended to the stimuli. These modifications were conducted bearing in mind the rules for creation of such stimuli suggested by Luh (1922; cited by Hilgard, 1951). The Hilgard CVCs used as base stimuli have a 53% association with real English words. All pseudo-words were checked against the Merriam-Webster Dictionary (2002) to ensure that no such words exist. Five stimuli were replaced as a result of this verification procedure.

Four-letter words were randomly selected from the Merriam-Webster Dictionary. Those words that had repeated letters were excluded from the set. In total, 160 unique experimental stimuli were developed. An additional 20 were also created for the purposes of training, 5 for each letter condition.

Apparatus

The experiment was designed and presented with Cedrus SuperLab Pro Version 2.0 on a 17" monitor operating at a resolution of 1024x768 pixels at 85Hz. As this experimental software package limited the response input to one character, an answer sheet was developed with 160 response fields, each with 8 positions where the participants were to identify the appropriate stimuli.

Procedure

Participants were tested individually in a dimly lit room. Participants were told that an eight-item sequence of four letters and four numbers would be presented, with the numbers random and the four letters conditions. Each participant was provided with an answer sheet and instructed to write down the memory stimuli after a response cue indicating the direction of recall, either “FORWARD’ or “BACKWARD”. The response cue followed a 250 ms inter-stimulus interval that occurred after the stimuli were presented. If the cue indicated forward, participants would write down the stimulus in the format it was presented in. Alternatively, if the cue indicated backward, participants would be required to respond with the order of the categories switched while preserving the order of the items within the category. Instructions emphasized speed and accuracy equally.

Twenty training trials were presented prior to the experimental blocks. The stimuli were presented in an identical manner to the experimental stimuli. This was done to ensure that the participants were familiar with the procedures and the type of stimuli being used.

The participants then performed two experimental blocks of trials, each of which used all 160 unique stimuli. Participants were randomly assigned to a distractor or no-distractor condition first, after which they received the alternate condition. During one block, a ‘+’ would appear in a random quadrant 250 ms before the memory stimulus as a proactive attention distractor. The same stimuli were presented during distractor and no distractor blocks with the opposite recall cue used in one block relative to the other. Stimuli were presented for 750 ms. Trials would occur randomly in accord with SuperLab 2.0 randomization controls. This resulted in 320 experimental trials, evenly distributed between the word categories.

Scoring

Participants were scored according to an all-or-none scoring of individual items when a correct item appeared in the correct order. In order to compare our current experiment with those of previous studies (e.g., Greene & Lasek, 1994) we first aggregated the score for each position. If a correct item appeared in the correct position as the memory stimulus, it was awarded 1.0. In the aggregated position analyses, all positions were summed together for a maximum of 8.0 points.

Results

Aggregated Position Analyses

A repeated-measures ANOVA was conducted using accuracy as the dependent variable. The factors that were examined included letter category (random letters, rhyming letters, pseudo-words, and words), category order (letters first or second), recall order (forward or backward), and score component (numbers or letters). Although the analysis treats the data as though there are only 4 categories constituted of letters, it is beneficial to observe that there are five categories with the inclusion of numeric stimuli. However, since the numeric stimuli are present in every display, they are treated as a separate variable. In addition, the distractor was also used as a within-subjects variable to replicate Greene and Lasek’s (1994) Experiment 3.

Our data was first analyzed as an aggregate score, which did not consider either the individual positions or score component. This was done in order to create a direct comparison between the present study and previous research. The following analyses present results in Greenhouse-Geisser adjusted values. Only those results involving the main effects or interactions involving category-order are reported.

In the present analyses, the main effect of category order demonstrated significance, $F(1, 22) = 22.828$, $MSE = .317$, $p < .001$, $\eta^2_p = .509$. The mean recall score demonstrated an improvement in performance when letters preceded numbers, rather than the reverse order. This main effect was qualified by the interaction category order and letter category, $F(3, 66) = 12.806$, $MSE = .230$, $p < .001$, $\eta^2_p = .368$. In general, pseudo-words or words were recalled with greater accuracy compared to random or rhyming tetragrams, but this difference was substantially greater when they preceded number tetragrams compared to when they were presented afterward. An additional interaction of category order and score component was also significant $F(1, 22) = 8.570$, $MSE = 13.702$, $p < .01$, $\eta^2_p = .280$. This interaction demonstrates that recall performance was substantially greater for letter stimuli when presented first ($M = 3.22$), then when presented second ($M = 2.52$) but that recall performance for numeric remained relatively constant over this manipulation ($M = 2.0$ and $M = 2.43$, respectively). Lastly, the above interactions were qualified by a significant 3-way interaction between category order, letter tetragram and score component, $F(3, 66) = 7.728$, $MSE = 2.535$, $p < .005$, $\eta^2_p = .260$. This finding demonstrates that although the category order effect is essentially additive when examining recall performance for letters, when we examine recall on the numeric tetragrams, performance increases only when they preceded random or rhyming letters. Lastly, although distractor did not have an overall effect on recall, it did significantly interact with category order and score component, $F(1, 22) = 4.549$, $MSE = 1.010$, $p <.05$, $\eta^2_p = .171$. The performance gains constituting the category-order effect were somewhat reduced when a distractor was presented before the stimuli. This seems to indicate reduced attention for the tetragram in the first position due to resources allocated to the distractor.

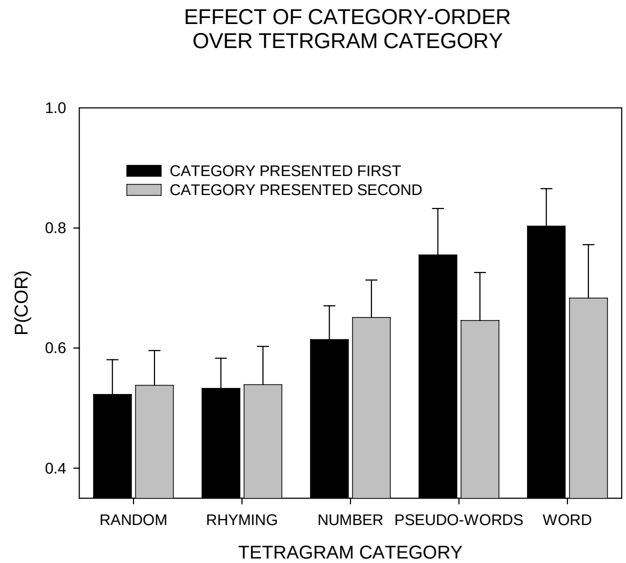
Serial-Order Analyses

A second analysis was performed on the same data with two important differences. First, numeric stimuli were included with letter stimuli to examine the effect of each category of tetragram (random letters, rhyming letters, pseudo-words, words and numbers). Second, each position was coded separately to obtain a fine-grained analysis of category-order effects on the positions. This was done to examine whether some or all of the positions in a display enjoy performance improvements.

In that it was based on the same data as the analysis reported above, it has many of the same features. After reviewing the similarities we will then examine the differences in these results.

The category-order effect was again significant but was subject to a decrease in effect size in comparison to the initial analysis, $F(1, 22) = 19.195$, $MSE = .116$, $p < .001$, $\eta^2_p = .466$. Category-order also interacted with tetragram category, $F(4, 88) = 8.588$, $MSE = .122$, $p < .001$, $\eta^2_p = .281$. As the mean recall performance in Figure 1 demonstrates, both random and rhyming letters as well as numeric tetragrams were unaffected by the order of category presentation, whereas pseudo-word and words exhibited reduced accuracy when presented second. Error bars represent Bonferroni significance values.

When position was included in the analysis it proved to be significant, $F(3, 66) = 66.113$, $MSE = .088$, $p < .001$, $\eta^2_p = .750$. In line with the literature of serial-position effect, item recall exhibited both primacy and recency effects. This effect was qualified by an interaction with category-order, $F(3, 66) = 4.842$, $MSE = .023$, $p < .05$, $\eta^2_p = .180$. When letter stimuli were presented first there was an overall increase in performance relative to when numbers were presented first. Additionally, when number tetragrams were presented first recency effects decreased making the final positions almost equal.



In this analysis, distractor did not have any significant effects alone or interact with category order.

Discussion

This experiment demonstrated that recall performance is improved when more salient stimuli precede less salient stimuli. Unlike previous studies, it shows that numeric stimuli do not enjoy primacy over all stimuli. When placed on a par with randomly generated letter stimuli, numbers did demonstrate relatively higher recall.

More generally, the results of Experiment 1 indicate that the category-order effect is the result of a complex interplay of cognitive mechanisms. The distractor manipulation only interfered when we considered each score component. When examining over position, the effect of distractor resulted in decreased performance when it was presented prior to a salient stimulus.

Experiment 2

Experiment 2 was designed to offer a further level of commensurability between past findings of the category-order effect. Since this was not a full replication of either Brooks & Watkins (1990) or Greene & Lasek (1994) it may be that some elements are missing which lead to their results. One element that was conspicuously absent in Experiment 1 was *successive* presentation of stimuli. It may be that unitary presentation causes greater primacy effects while temporally sequential presentation leads to greater recency effects. To this end, stimuli were broken up into their constituent score components, which were presented one after another. Presentation rate was also sped up to limit the encoding advantage of having fewer digits simultaneously presented.

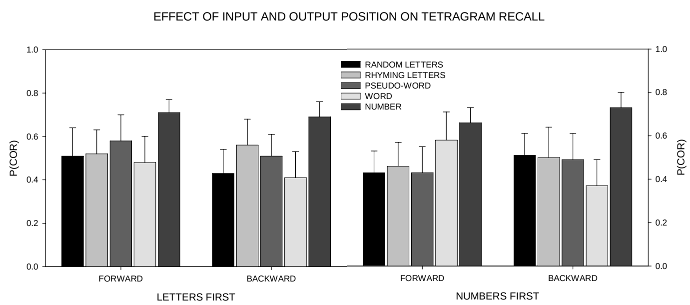
Method

Participants

Twenty-one undergraduates (13 females, 6 males) participated, and were awarded 1% toward their final grade in an introductory psychology class.

Materials and Procedures

The present experiment segmented the identical stimuli from Experiment 1 by score component, and presented them in rapid succession (i.e. present BONE then 1234) at a rate of 250ms/score component. Procedures and apparatus were the same as in Experiment 1.



Results

Serial-Order Analyses

An ANOVA with the same dependent and independent variables was used in the current experiment as in Experiment 1. Instead of an aggregate analysis, the present experiment confined itself to an examination over serial-order.

Replicating the serial-order analysis of experiment 1, the effect of category-order was significant, $F(1, 18) = 6.137$, $MSE = .038$, $p < .05$, $\eta^2_p = .254$. When letters preceded numbers recall improved ($M = 0.54$) compared to when numbers preceded letters ($M = 0.52$). The interaction of category-order and recall order was also significant, $F(1, 18) = 13.441$, $MSE = .019$, $p < .005$, $\eta^2_p = .427$. Participants' average recall was uniformly higher when they were required to report letters first followed by number stimuli. This effect was significant larger when words were presented first than when they were presented second. The interaction of category-order and tetragram category was also significant, $F(4, 72) = 4.529$, $MSE = .047$, $p < .01$, $\eta^2_p = .201$. Interestingly, recall for numeric tetragrams was uniformly greater than any of the letter tetragrams. When rhyming letters or pseudo-words were presented first, they were recalled with greater accuracy than rhyming letters or words. When number stimuli were presented first, letter stimuli were recalled at an equivalent rate. Qualifying all of these interactions was the significant three-way interaction of category-order, recall order and tetragram, $F(4, 72) = 8.526$, $MSE = .065$, $p < .005$, $\eta^2_p = .321$. As Figure 3 demonstrates, a category-order effect was present for numbers when they were presented first but recalled second, for random and rhyming letters as well as pseudo-words when presented second but recalled first and for words when presented first and recalled second. As in experiment 1, error bars represent Bonferroni significance values. The result again demonstrates the robustness of the category-order effect while also providing further evidence for interference that can result from rapid presentation.

Discussion

Experiment 2 demonstrated improved recall performance when letters were presented first even under conditions that otherwise favoured the recall of number tetragrams. When stimuli were presented quickly, recall for numeric stimuli was significantly greater than all letter stimuli. However, when there was a category-order effect it favoured recall performance for letter stimuli.

Interestingly the category-order effect for letter tetragrams occurred within *different* recall orders, indicating that the use of item information varies depending on the temporal order that a stimulus is presented and retrieved in. The rapid presentation rate is no doubt the driving factor behind the reduced recall for word stimuli, but this effect needs to be further qualified as it only occurs when letters are presented and recalled in the second position.

General Discussion

The present research replicated a category-order effect similar to that observed by Brooks and Watkins (1990) and Greene and Lasek (1994). Previous research demonstrated that when numbers preceded words, recall performance on the entire list was improved. In contrast, in Experiment 1 we found that while a category-order effect is indeed evident within the overall score, it only exhibits improved performance for word and word-like stimuli, with no score improvement with random letters, rhyming letters, or random numbers. By increasing stimulus presentation rate and mode of presentation in Experiment 2, we again observed a category-order effect but one in which rhyming letters now exhibited the effect under certain conditions. Rapid presentation also nearly eliminated the category-order effect in word stimuli. This supports early work that presentation rate can nearly erase the category-order effect (Greene & Lasek, 1994).

The difference between the present studies and other research on the category-order effect are instructive. In Experiment 1, each stimulus was presented as a single string, which would most likely facilitate grouping of items in memory at the time of encoding. This should consequently reduce demands on the working memory system resulting in improved performance for the overall scores. A final contribution of Experiment 1 was evident when position was included as a dependent variable; demonstrating that a serial-order analysis is an effective means to study the category-order effect.

In Experiment 2, the letter and number tetragrams were presented in rapid succession instead of a single string. Consequently, participants have little time to examine the stimuli to encode properties that could be used to access comparable information in long-term memory. Even though we did not observe a strong category-order effect for number tetragrams, it appears that successive presentation favours encoding and retrieval of numeric information when it is placed in competition with letter information. This performance gain is offset by improvements resulting from the category-order effect for letter stimuli

Further study is required to bridge the gap between the present study and past research on the category-order effect. Taken together with other studies, the present experiments suggest four key findings. First, the category-order effect is a robust phenomena that is evident in a number of stimuli sets over diverse methods of presentation. Second, this effect exists at both a macro-level (aggregate scores for lists) and a micro-level (positional effects). Third, the category-order effect is *relative* to the two categories under consideration. Last, and perhaps most important, the category-order effect is not limited to instances when a small homogenous category is presented before a larger heterogeneous category. Instead, knowledge stored in long-term memory and the presentation mode can facilitate the encoding and retrieval of information in working memory.

One limitation of this study should be noted. This research only controlled for the lexicality of the letter stimuli while preventing the number stimuli from being facilitated by any extrinsic knowledge or intrinsic features. A possible direction for future research would be to use historical dates (e.g., 1939), numbers with regular increases in magnitude (e.g., 1234 or 3579), or the binary sets / repeating sets (e.g., 1011).

In short, our findings suggest that the category-order effect is a result of a complex interaction of numerous well-established properties of memory. When stimuli can be grouped together on the basis of categorical properties (e.g., lexicality) stored in long-term memory, recall on a list is improved when they are presented prior to stimuli from a category with knowledge that is less accessible. In this sense, the category-order effect can be understood as a significant measure of the interaction of cognitive processes. The present study highlights the need to analysis these phenomena at an appropriate level to ensure that the contributions of other memory process are taken into account.

Acknowledgements

The authors would like to thank Sylvain Pronovost and Michael Henighan for their aid in coding the experimental results and past collaborative efforts.

References

Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J.T. Spence (Eds.), *The Psychology of Learning and Motivation (Vol. 2)*. London: Academic Press.

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

Brooks, J. O., & Watkins, M. J. (1990). Further evidence of the intricacy of Memory Span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 1134-1141.

Feldman, J. (2003). The simplicity principle in human concept learning. *Current Directions in Psychological Science,* 12(6), 227-232.

Greene & Lasek (1994). Category-order effects in memory span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1391-1395.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, 13, 644-645.

Hilgard, E. R. (1951). Methods and procedures in the study of learning. In S. S. Stevens, *Handbook of experimental psychology*. New York: Wiley.

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

International Phonetic Association (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. England: Cambridge University Press.

Kausler, D. H., & Puckett, J. M. (1979). Effects of word frequency on adult age differences in word memory span. *Experimental Aging Research*, 5, 161-169.

Landauer, T. K., & Freedman, J. L. (1968). Information retrieval from long-term memory: category size and recognition time. *Journal of Verbal Leaning Behaviour*, 7, 291-295.

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

Miller, G. A. (1956/1994). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 101, 343–352.

Murdock, B. B., Jr. and von Saal, W. (1967). Transpositions in short term memory. Journal of Experimental Psychology, 74, 137-143.

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

Reichler, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275-280.

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

Waugh, N. C., & Norman, D. A. (1965). Primary Memory. *Psychological Review*, 72, 89-104.