UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Learning and failing to learn in immediate memory

Permalink
https://escholarship.org/uc/item/08k132fh

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Authors
Beaman, Philip
Roeer, Jan

Publication Date
2009

Peer reviewed

Learning and failing to learn in immediate memory

C. Philip Beaman (c.p.beaman@reading.ac.uk)
Department of Psychology, University of Reading
Earley Gate, Whiteknights, Reading RG6 6AL, UK

Jan P. Röer (Jan.Roeer@uni-duesseldorf.de)
Department of Experimental Psychology, Heinrich-Heine-Universität
Düsseldorf, Germany

Abstract

Two experiments examine the effect of regular feedback on the appearance of the word-length and the irrelevant sound effects in immediate memory. A reliable effect of learning was observed but both effects persisted across multiple learning trials, contrary to suggestions that they could be diminished either by feedback-inspired strategy change or by dual-task learning. Learning improved performance most noticeably at the end of the serial position curve but there was no sign that the irrelevant speech eliminated the word-length effect at any stage of learning, consistent with Tremblay et al. (2000) but inconsistent with Neath et al. (1998). Implications for immediate memory effects, and for models of intra-individual variation in immediate memory, are considered.

Keywords: Immediate ("working") memory; irrelevant speech; word-length.

Introduction

The relationship between immediate recall and long-term learning is of long-standing interest in cognitive psychology (Thorn & Page, 2009). The structure of immediate memory (sometimes known as "working memory"), at least in the verbal domain, is defined by a number of well-documented effects which are presumed to uniquely identify the involvement of particular mechanisms such as the phonological loop (Baddeley, 1986) in any immediate or working memory task. Two such effects are the word-length effect and the irrelevant sound effect.

The *word-length effect* (Baddeley, Thomson & Buchanan, 1975) refers to the finding that immediate recall of longer words is worse than immediate recall of similarly sized lists of shorter words. The dominant explanation of the effect is that immediate memory is restricted by the time it takes to rehearse by articulation a to-be-recalled list (Baddeley et al., 1975). The *irrelevant sound effect* (Beaman & Jones, 1997) refers to the finding that, in the presence of noise (usually speech), immediate memory performance is reduced. To produce this result, the speech or sound usually has to *change-in-state* or, in other words, abruptly alter in pitch or other characteristics. One account states that changes in state within the irrelevant sound perturb episodic rehearsal of the to-be-recalled material (Jones & Macken, 1993) whilst other widely-held theories suggest an involuntary division of attentional resources between to-be-recalled and irrelevant material. Both effects can be mediated by instruction in strategy. For example, the word-length effect

is reduced or reversed if participants are asked to think of semantic links between words prior to presentation of the words (Campoy & Baddeley, 2008) and, similarly, the effect of irrelevant sound is eliminated when instructions are given to recall by category membership rather than in serial order (Perham, Banbury & Jones, 2007). The two effects are also related in the disputed claim that irrelevant speech eliminates the word-length effect (Neath, Surprenant & LeCompte, 1998; Tremblay, Macken & Jones, 2000).

It is not clear, however, whether individuals can spontaneously, or as a result of learning, avoid the detrimental effects associated with short-term storage (such as those of word-length or irrelevant sound) by shifting away from a phonological loop-type structure onto some other form of memory. This question is of particular interest because failure to find such effects under particularly exacting memory loads has been taken as an indication that the encoding of those memory loads has been strategically-shifted away from phonological storage onto other forms of memory and hence the results of these experiments have been taken as not applicable to the phonological loop construct (Baddeley, 2000; Baddeley & Larsen, 2003, 2007; Campoy & Baddeley, 2008; Larsen & Baddeley, 2003). If there is no evidence for learned strategies that bypass irrelevant sound/word-length effects then the results of these experiments will need to be reinterpreted and their implications reconsidered.

It is also not clear in what ways the basic pattern of performance might be affected by repeated practice and learning. A general improvement due to practice might be anticipated but it is not obvious what is learned and how this might manifest itself (e.g., will the improvement be centered on the earlier items and hence reflect some kind of improvement in rehearsal/ consolidation?) Nor, finally, is it clear how the word-length and irrelevant sound effects might interact with each other and with learning and practice. The account of irrelevant sound disruption provided by the feature model for example (Neath, 2000), involves two separate processes. Firstly, it suggests that the effect of changes-in-state should be considered as a dual-task cost incurred by dividing attention between relevant and irrelevant streams. As such, one would also expect a reduction in the size of the effect if, but only if, opportunity is given to learn to dual-task effectively. To date, appropriate studies have not been carried out to test the dual-task cost explanation. Passive listening to the speech in

the absence of an immediate memory task (Banbury & Berry, 1998; Jones, Macken & Mosdell, 1997) is inappropriate to test for the ability to dual-task. Some studies have combined the speech with a memory task in a manner more conducive to improving dual-task capabilities (e.g., Hellbrück, Kuwano & Namba, 1996) but none have provided participants with feedback on their performance under irrelevant sound conditions. Since many individuals do not realize that they are performing poorly in the presence of distracting sound (Beaman, 2005; Ellermeier & Zimmer, 1997) there is limited opportunity to learn dual-tasking skills, or change in strategy, that might protect against the irrelevant sound effect.

Running irrelevant sound and word-length manipulations simultaneously also gives the opportunity to examine the second process invoked by the feature model in accounting for the irrelevant speech effect. In four experiments, Neath et al (1998) found that irrelevant speech, but not irrelevant sound, eliminated or masked the word-length effect. This is predicted by the feature model because, within that model, the attentional effects of changing-state irrelevant sound already outlined are accompanied by the adoption of elements of the speech into the representations of to-be-recalled items, a process known as feature adoption, masking the word-length effect (Neath et al., 1998, Neath, 2000). Against this, Tremblay et al. (2000) found no evidence that the word-length effect was masked in the manner suggested and hence rejected the feature model account. Whilst there is no suggestion that any word-length/irrelevant speech interaction should necessarily be affected by learning, a repeated measures experiment in which word-length and irrelevant sound are tested on multiple occasions should provide strong evidence to support one view or another.

Experiment One

The first experiment examines the long-term effect of word length and irrelevant non-speech sounds over multiple testing sessions. No word-length by irrelevant sound interaction is predicted for this experiment, rather we hope simply to document any changes in either or both effects, together with the overall patterns of improvement expected as a result of learning and feedback.

Method

Participants. Eighteen undergraduate students of the University of Reading, (one male seventeen female), participated in return for course credit. All reported normal hearing and normal, or corrected to normal, vision.

Materials and Design. Stimuli were presented using E-Prime experiment generation software. The to-be-remembered stimuli were 120 long (three- to five-syllable) and 120 short (single-syllable) words taken from LaPointe and Engle (1990). For each block 24 words were used per word-length condition, randomly sampled to provide six to-be-remembered words per trial. Different sets of 24 words were used for each block. There were 40 trials per block, 10 in each condition, and 5 blocks in total. Word length was blocked and alternated between participants, irrelevant sound was randomized separately for each participant. One trial consisted of six words presented in the centre of the screen in 24 point Arial font. The words were presented consecutively at a rate of 1/ s, 400ms on, 600 ms off. After the final word, there was a 2s pause before the presentation of a 4 x 6 matrix of words, presenting the participant with the six target words together with eighteen lures, arranged in alphabetical order from left to right within the matrix. A series of six numbered "responses boxes" were also presented at the right-hand side of the screen. All words, targets and lures, were presented in 24 point Arial font. The irrelevant sound stimuli consisted of six square wave tones, of frequency 440, 880, 1240, 620, 1080 and 360 KHz. Each tone lasted for 400ms and was presented concurrently with visual presentation of a to-be remembered word. The first two tones of the sequence were then played once more during the 2s retention pause at the same presentation rate. The presentation order of the tones was identical for each participant and did not alter over the experiment. All testing was carried out in a sound-attenuated cubicle.

Procedure. Participants were informed that we were interested in how accurately they could remember the order in which a series of words were presented, and how well they could learn to remember the order and ignore any extraneous background sound. Each participant completed 5 blocks of 40 trials, with blocks being completed on consecutive weekdays. Participants were told that, following presentation of each trial, they should "drag and drop" the words they saw, in the order in which they saw them, into the correct response box. For example, if they thought the word *break* was the first word they saw, they should click on this word and drag it to the response box 1 (the top-most box) and then release. Null responses, recording a failure to recall, were made by leaving a response box blank. Following each trial participants' performance was fed back to them by coloring the boxes containing correct responses in green, and those containing incorrect responses in red. There were 3 breaks, one after every 10 trials to allow participants to rest. Breaks were self-terminated by the participant and lasted up to 1 minute. Each testing session lasted approximately 60 minutes.

Results.

Two participants were dropped from the analysis for failure to complete the five blocks. Repeated measures analysis of variance (ANOVA) on number correct shows a main effect of word-length, $F(1, 15) = 184.02$, $p < .001$, and a main effect of learning, $F(4, 60) = 9.85$, $p < .001$, but no effects of irrelevant sound, $F < 1$. None of the interactions reached significance (all ps $> .05$). The effect of learning on performance in both word-length conditions is shown in Figure 1.
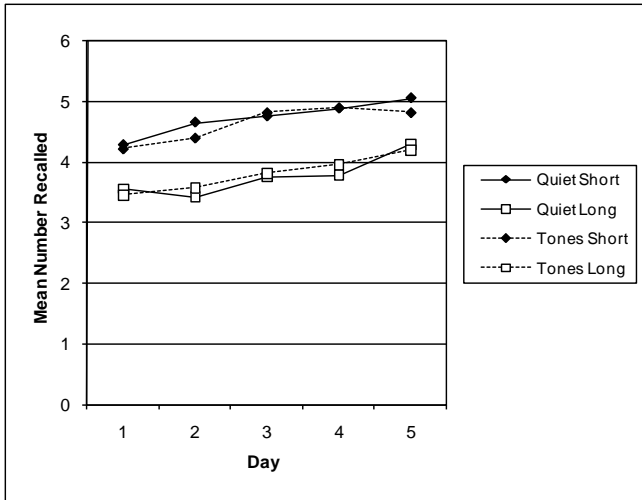
Figure 1: The effects of learning on recall in two word-length conditions (long, short) and two noise conditions (tones, quiet)

Serial position functions were also compared between the first and final day. As shown in Figure 2, these display an upward shift in performance over time restricted to the final four serial positions.
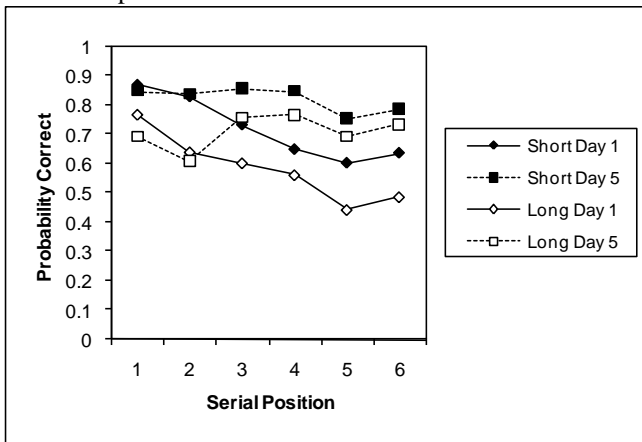

Figure 2: Serial position curve for first and final testing blocks in short and long word conditions.

Discussion

The lack of an irrelevant sound effect with the tone stimuli limits the conclusions that can be drawn regarding this effect. Tones are known to be quantitatively less disruptive than speech whilst showing qualitatively similar performance (Jones & Macken, 1993), so the null effect here is uninformative as it most likely the result of a weak experimental manipulation (e.g., low "dose" of relatively unchanging stimuli). Of interest, however, is the finding that the appearance of the word-length effect is spared across learning trials. The upward shift in performance across the final four serial positions as a consequence of learning is

also of interest as it suggests a possible means of accounting for the learning effect within the feature model.

As noted by Beaman, Neath and Surprenant (2008), there are limited ways in which performance can be varied within the feature model without assuming some variation in the stimuli themselves. When variations in the stimuli are ruled out, only two free parameters in the basic model (Nairne, 1990) affect overall performance. Restrictions on space preclude more detail explication here, but Beaman et al. (2008) explored the possibility that variation in one of these parameters might provide an appropriate basis for modeling inter-individual differences in immediate memory using the feature model. This left the basis of intra-individual differences unspecified, relying upon noise in the system to provide trial-by-trial differences in performance (Beaman et al., 2008). By a process of elimination, however, if intra-individual variation is in any way systematic and not subject only to random noise only one free parameter in the feature model can possibly be employed to reflect this. The effect of manipulating the remaining free parameter, r, which models the number of retrieval attempts, is to flatten the end of the serial position curve whilst leaving the word-length effect intact (see Figure 3). This is similar in action to the learning effect observed in Experiment 1. Tentatively, therefore, it seems as though an appropriate learning parameter might be available in this model. Figure 3 therefore represents the feature model's prediction for the general shape of the serial position curves following learning. As previously established by Neath et al. (2000), the model also predicts that the word length effect preserved under learning with feedback will be abolished by the presence of irrelevant speech and Experiment 2 examines this prediction.
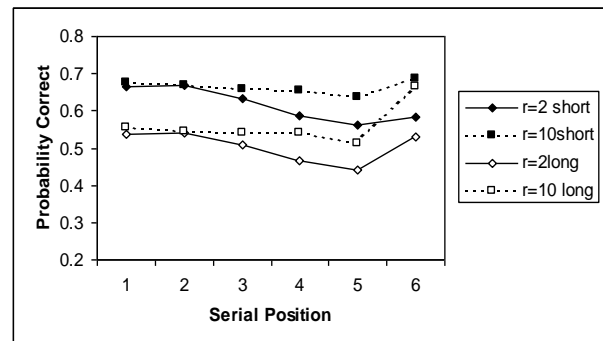

Figure 3: Learning in the feature model as a function of increased retrieval attempts. Variation in the retrieval parameter r reflects the extent of learning on recall of short- and long-word lists.

Experiment Two

Experiment 1 provided useful evidence regarding the word-length and learning effects but failed to show an irrelevant sound effect, and is therefore uninformative regarding one of the major reasons for running the study. Experiment 2 rectifies this by examining the more potent irrelevant *speech* effect under similar circumstances. This

will provide a further check on the main findings of Experiment 1 and will also allow for examination of whether the effect is, wholly or in part, reduced by dual-task learning and whether the word-length effect is affected by the presence of irrelevant speech at any point. This latter investigation will provide confirmatory evidence either for Neath et al. (1998) or Tremblay et al. (2000).

Method

Participants. Eighteen undergraduate students of the University of Reading, (three male, fifteen female), participated in return for course credit. None had taken part in Experiment 1. All reported normal hearing and normal, or corrected to normal, vision.

Materials and Design. To-be-remembered stimuli were as in Experiment 1. Design was also as in that experiment with the exception that, for logistic reasons, only four consecutive testing sessions were possible. The irrelevant speech stimuli consisted of the words London, Purple, Swindon, Yellow, Cardiff, Crimson spoken in a male voice recorded at 22.05KHz and played back at the same rate as in the previous experiment. As previously, the first two items of the sequence (London, Purple) were played once more during the 2s retention pause. The presentation order of the words was identical for each participant and did not alter over the experiment.

Procedure. Procedure was as Experiment 1.

Results

Three participants were dropped from the analysis for failure to complete the four blocks. Repeated measure ANOVA on the remainder revealed a main effect of word-length on accurate recall, $F(1,14) = 52.34$, $p < .001$, a main effect of speech, $F(1, 14) = 38.79$, $p < .001$, and a main effect of learning, $F(3, 42) = 4.3$, $p = .024$ (Huynh-Feldt corrected for non-sphericity). There was no interaction between word-length and speech, $F < 1$, consistent with Tremblay et al. (2000), and no significant interactions involving word-length and learning, $F(3, 42) = 1.82$, $p = .16$, speech and learning $F(3,42) = 1.67$, $p = .2$ (Huynh-Feldt corrected for non-sphericity), or word-length, speech and learning $F(3, 42) = 1.42$, $p = .25$. The effect of learning on all four conditions is shown in Figure 4. This figure shows an upward trend for performance over time in all conditions.
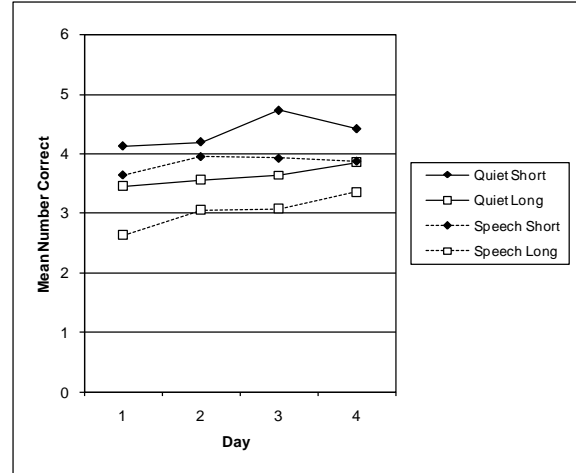

Figure 4: The effects of learning on recall in two word-length conditions (long, short) and two noise conditions (speech, quiet).

The serial position functions on days 1 and 4 are given in Figures 5a and 5b. These figures show detrimental effects of word-length and irrelevant speech on performance at all serial positions during the first (Figure 5a) and final day (Figure 5b), and the effects of learning on the serial position curve. As in the previous experiment, the effect of learning is to flatten the bow-shape of the serial position curve (Figure 5b) from its initial form in block 1 (Figure 5a).
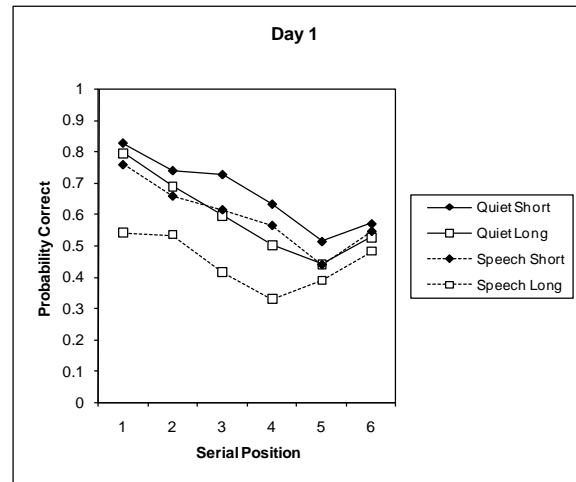

Figure 5a: Serial position curve for the first testing block (day 1) in short and long word conditions.
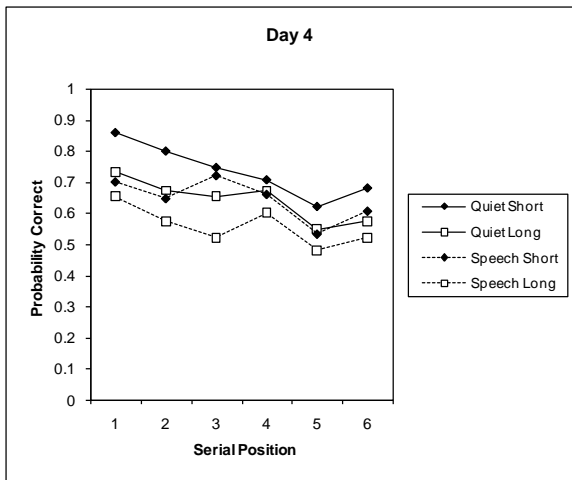
Figure 5b: Serial position curve for the final testing block (day 4) in short and long word conditions.

Discussion

Experiment 2 replicates the main finding of Experiment 1 that the word-length effect is robust over multiple learning episodes when feedback is given. Additionally, Experiment 2 establishes that the effects of irrelevant speech are likewise robust. Multiple learning episodes, even when direct and immediate feedback is given, do not eliminate the impact of irrelevant speech on immediate memory. It seems unlikely, therefore, that the irrelevant speech effect is, wholly or in part, the result of dividing attention where the participant is unaware that divided attention results in poorer performance on the main (immediate memory) task. This is contrary to the predictions of the feature model (Neath, 2000). A third observation from Experiment 2 is that the word-length effect also remained robust across multiple learning trials in the presence of irrelevant speech. This result is consistent with the report by Tremblay et al. (2000) that irrelevant speech does not eliminate the word-length effect. It is, however, inconsistent with the original report of such an observation by Neath et al. (1998) and casts doubt upon the feature model's account of the irrelevant speech effect. Finally, the experiment shows that – with the exception of the long words plus speech condition, where performance was surprisingly poor at primacy in Block 1 – the effect of learning is largely restricted to the final four serial positions, as in Experiment 1. This latter result is consistent with the account summarized in Figure 3, in which learning within the feature model was linked to increases in a retrieval parameter.

General Discussion

The two experiments reported here have shown an effect of learning-to-learn that preserves two key phenomena across learning trials – the word-length and irrelevant speech effects. These phenomena are preserved despite their openness to instructed strategy shifts (e.g., towards semantic-based strategies; Campoy & Baddeley, 2008; Perham et al., 2007), suggesting both that feedback-driven learning need not involve any change in strategy to be effective, and that information on performance levels is insufficient by itself to bring about strategy change. This latter point casts doubt on suggestions that changes in the the appearance of immediate memory phenomena result from switching to semantic memory when performance levels are perceived to be inadequate using phonological encoding (Baddeley & Larsen, 2007). In the current experiments, participants were made aware, by means of direct and immediate feedback, that lists of long words and lists accompanied by irrelevant speech result in relatively poor performance. Despite this, there is no sign that the participants spontaneously shifted strategy to improve performance in these conditions even though, as already documented, instructions to use particular strategies can change the appearance of both these effects. Although no strategy shift was evident in these experiments, performance did generally improve, however.

As regards the irrelevant speech effect, results show preserved levels of irrelevant speech interference across multiple learning trials with feedback. This lends no support to the idea that divided attention contributes to the size of the effect since one would expect that practice in dividing attention, coupled with feedback indicating the extent to which irrelevant speech is affecting performance on the primary task, would reduce any dual-task cost and hence the interference observed. Similarly, the persistence across multiple learning episodes of the word-length effect in the presence of irrelevant speech casts further doubt on the feature adoption explanation of the irrelevant speech effect. In sum, these experiments provide no support for either part of the feature model's account of irrelevant speech effects. The absence of a word-length by speech interaction is also inconsistent with suggestions that changing-state speech and changing-state concurrent articulation ("articulatory suppression") are functionally equivalent (Macken & Jones, 1995), as the latter is reliably seen to mask the word-length effect (e.g., Baddeley, Lewis & Vallar, 1984).

One aspect of the feature model that receives more support from the current findings is the suggestion that a single parameter, r, could reflect learning rate within the model. This would allow for a source of intra-individual differences other than random noise. At present, most models of immediate memory lack a principled account of either inter-individual or intra-individual differences. It would be no small achievement to unambiguously identify such individual difference factors and care must be taken not to overstate the present case.

In the first instance, it could be argued that the serial position curve flattens with learning simply because the primacy portion of the curve (the first two serial positions) is already at ceiling. Against this, it is obvious from Figures 2 and 4 that performance at these two serial positions is not at ceiling in any of the long word or irrelevant speech

conditions. In 4 of the 8 serial positions concerned in Experiment 2 performance was numerically worse after learning than before, as it was for 3 of the 4 serial positions considered in Experiment 1. A second and more substantial point is that merely identifying a parameter in the model which can be varied to alter performance in the appropriate way does not necessarily provide a satisfactory explanation of the learning effect. The r parameter might plausibly be regarded as measure of motivation rather than retrieval per se but, given that participants received direct feedback on their performance in the current experiments, the two are to some extent conflated in the present study.

Regardless of the merits of these arguments, the current study demonstrates both strengths and weaknesses of the feature model when opportunity for learning is given. Importantly, the study also demonstrates that word-length and irrelevant speech effects are stable over both time and feedback and, as such, neither effect seems open to spontaneous strategy-shift although both are affected by instruction (Campoy & Baddeley, 2008; Perham et al., 2007). Finally, the study reinforces the conclusions of Tremblay et al. (2000) in finding no masking of the word-length effect by irrelevant speech suggesting that, in this regard at least, irrelevant speech and concurrent articulation are functionally distinct

Acknowledgments

References

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

Baddeley, A. D. (2000). The phonological loop and the irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review, 7*, 544-549.

Baddeley, A. D., & Larsen, J. D. (2003). The disruption of STM: A response to our commentators. *Quarterly Journal of Experimental Psychology, 56A*, 1301-1306.

Baddeley, A. D., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *Quarterly Journal of Experimental Psychology, 60*, 497-504.

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

Banbury, S. & Berry, D. C. (1998). Habituation and dishabituation to speech and office noise. *Journal of Experimental Psychology: Applied, 3*, 1-16.

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

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

Beaman, C. P., Neath, I., & Surprenant, A. M. (2008). Modeling distributions of immediate memory effects: No strategies needed? *Journal of Experimental Psychology: Learning, Memory & Cognition, 34*, 219-229.

Campoy, G., & Baddeley, A. (2008). Phonological and semantic strategies in immediate serial recall. *Memory, 16*, 329-340.

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

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

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological encoding in working memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 21,* 103-115.

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

LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory & Cognition, 16*, 1118-1133

Larsen, J. D., & Baddeley, A. D. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression and manual tapping: Do they have a common source? *Quarterly Journal of Experimental Psychology, 56A*, 1249-1268.

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner ear: Single or double agency? *Journal of Experimental Psychology: Learning, Memory & Cognition, 21*, 436-448.

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

Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review, 7*, 403-423.

Neath, I., Surprenant, A. M., & LeCompte, D. C. (1998). Irrelevant speech eliminates the word length effect. *Memory & Cognition, 26*, 343-354.

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

Thorn, A., & Page, M. P. A. (2009). *Interactions between short-term and long-term memory in the verbal domain*. Hove: Psychology Press.

Tremblay, S., Macken, W. J., & Jones, D. M. (2000). Elimination of the word-length effect by irrelevant sound revisited. *Memory & Cognition, 28*, 841-846.