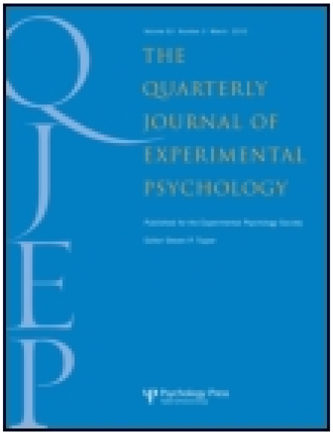

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative

Alan D. Baddeley [a] & Janet D. Larsen [b]
[a] York University , York, UK
[b] John Carroll University , University Heights, OH, USA
Published online: 13 Apr 2007.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2007, 60 (4), 497–504

Psychology Press
Taylor & Francis Group

Lead article

The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative

Alan D. Baddeley
York University, York, UK

Janet D. Larsen
John Carroll University, University Heights, OH, USA

Jones et al. (Jones, Hughes, & Macken, 2006; Jones, Macken, & Nicholls, 2004) identify the interaction between phonological similarity, articulatory suppression, and stimulus presentation mode in verbal short-term memory as potentially providing important support for the phonological loop hypothesis. They find such an interaction but attribute it to "perceptual organization masquerading as phonological storage". We present data using shorter letter sequences and find clear evidence of the interaction predicted by the phonological loop hypothesis, which, unlike the evidence of Jones et al., is not limited to recency, and which provides continued support for the phonological loop hypothesis.

Working memory is assumed to be a system that facilitates the temporary storage and manipulation of information and as such facilitates complex thought processes. The multicomponent model of working memory (Baddeley, 2000, in press; Baddeley & Hitch, 1974) assumes an attentionally limited controller, the central executive, a multidimensional temporary store, the episodic buffer, and two subsidiary storage systems, the visuo-spatial sketchpad and the phonological loop.

The phonological loop is assumed to have two principal components: a temporary store capable of holding information over a period of seconds and an articulatory rehearsal system capable of maintaining information by subvocalization and of converting visually presented but nameable items such as letters into a phonological code. This subsystem was initially termed the *articulatory loop*, but was renamed the *phonological loop*, on the grounds that the capacity for storage was the central feature of the system, which can operate without articulation, provided material is presented auditorily.

The first clues to the nature of the code involved in this temporary store came from

Correspondence should be addressed to Alan D. Baddeley, Department of Psychology, University of York, Heslington, York YO10 5DD, UK. E-mail: ab50@york.ac.uk

Thanks to Alicia January, Betsy Lehman, and Kristi Monday for their assistance with data collection and coding and to Dennis Norris and an anonymous referee for comments on an earlier draft. We are grateful for financial support from Medical Research Council Grant 001M8025701 to Alan Baddeley and Graham Hitch.

http://www.psypress.com/qjep DOI:10.1080/17470210601147572

Conrad's (1964) demonstration of the acoustic similarity effect, whereby errors in recalling sequences of visually presented letters tended to be similar in sound to the originals, and that sequences of similar letters were harder to recall than were sequences of dissimilar letters (Conrad & Hull, 1964). However, the fact that this effect occurred with visually presented items indicates that an acoustic input is not necessary and that the term "acoustic similarity" might therefore be misleading. Attempts to specify the nature of the code in more detail proved unsuccessful (Hintzman, 1967; Wickelgren, 1969) leading us to opt for an alternative term. We initially chose *phonemic*, but given that this had a rather precise linguistic meaning, opted instead for *phonological*, which we hoped would be relatively neutral with respect to the exact nature of the code, apart from indicating that it was capable of encoding the spoken features of language. The question of whether it also is capable of maintaining nonlinguistic information remains open. Salame and Baddeley (1989) observed that storage in this system can be disrupted by music but not white noise, a phenomenon that has been explored much more extensively in a series of subsequent papers by Jones and colleagues (See Jones, Beaman, & Macken, 1996, for a review). The exact nature of the code, however, remains an important but open question.

Evidence for the phonological loop hypothesis came from a range of converging sources, with the phonological similarity effect being assumed to be a marker of the store itself, while evidence for the process of rehearsal came both from the word length effect (Baddeley, Thomson, & Buchanan, 1975) and from the effect on immediate verbal memory of articulatory suppression, whereby the subject is prevented from rehearsal by the requirement to utter a repeated sound such as the word "the" (Murray, 1968). When material is presented visually, subvocalization is assumed to be necessary for registering the material in the phonological store, while access is assumed to be automatic for auditory material. Central to this view is the interaction between phonological similarity, articulatory suppression, and modality of presentation. When material is presented visually, suppression removes the phonological similarity effect, presumably because suppression prevents the phonological recoding of the visually presented letters. The similarity effect survives suppression, however, when presentation is auditory, consistent with the assumption that the stimuli gain direct access to the phonological store.

This crucial interaction has remained unchallenged until a series of recent ingenious experiments by Jones and colleagues (Jones, Hughes, & Macken, 2006; Jones, Macken, & Nicholls, 2004). Their initial study involved the auditory presentation of sequences of seven phonological similar or dissimilar letters, accompanied by silence or concurrent whispered articulatory suppression on the part of the subject. Jones et al. (2004) obtained the interaction predicted by the phonological loop hypothesis, but found that it was limited to the recency portion of the serial position curve (see Figure 1). They interpret this, very plausibly, in terms of the well-established modality effect whereby auditory presentation enhances the recency effect (Conrad & Hull, 1968; Crowder & Morton, 1969). Support for this view is provided by a subsequent experiment

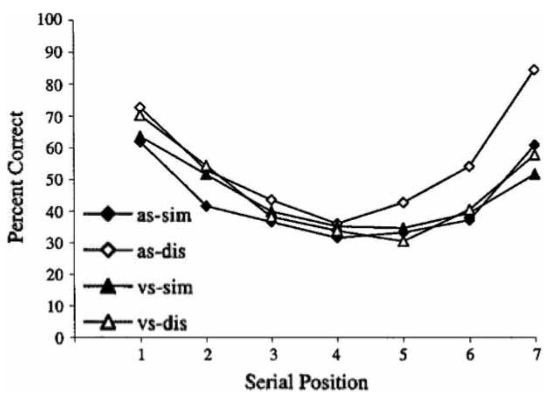

Figure 1. *The influence of phonological similarity and articulatory suppression on the immediate recall of consonant sequences. From "The phonological store of working memory: Is it phonological and is it a store?", by D. Jones, W. J. Macken, and A. P. Nicholls, 2004,* Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, *pp. 656–674. Copyright 2004 by the American Psychological Association. Reprinted with permission.*

in which a spoken suffix is found to remove the crucial interaction (Jones et al., 2004).

They point out that virtually none of the literature supporting the crucial interaction includes serial position curves; one exception is the study by Murray (1968) who reports the similarity effect throughout the list. However, this is discounted on the grounds that he uses atypical methodology. Jones et al. (2004) conclude that their data undermine the phonological loop hypothesis. They propose instead a hypothesis based on a combination of acoustic organizational processes and "gestural" rehearsal.

A possible limitation to the conclusions drawn by Jones et al. (2004) stems from the relationship between the phonological similarity effect and list length. As sequence length increases, a point is reached beyond which the phonological similarity effect tends to disappear (Baddeley, 1966a, 1966b; Baddeley & Larsen, 2003; Hall, Wilson, Humphreys, Tinzmann, & Bowyer, 1983; Hanley & Bakopoulou, 2003; Hanley & Broadbent, 1987; Johnston, 1982; Neath, Bireta, & Surprenant, 2003). A subsequent study by Jones et al. (2006) addresses this point by using sequences of only five letters, accompanied, however, by a suffix, and in their Experiment 3, a combined prefix and suffix. It is only in this latter study that the crucial interaction is totally confined to the recency portion of the serial position curve. Nevertheless, Jones et al. conclude that, "Under suppression, the survival of the phonological similarity effect with auditory presentation is due largely to the survival of performance for the last one or two phonologically dissimilar items, i.e. the recency portion of the serial position curve" (Jones et al., 2006, p. 267).

With a view to encouraging caution, we would like to present some of our own data that do not appear to support the claim that the crucial interaction is confined to the recency effect. The experiments in question form part of an extension to a series of studies by Larsen and Baddeley (2003) investigating the influence on immediate serial recall of phonological similarity, irrelevant speech, tapping, and articulatory suppression delivered at the different rhythms. This initial series of studies involved visual presentation, while our ongoing extension involves contrasting visual and auditory presentation, again studying the same range of variables. We again opted for sequences of six letters, sufficient to avoid ceiling effects while reducing the chance of pushing participants to a point at which they abandon phonological coding. We have in fact conducted three studies in which participants attempted to recall sequences of similar or dissimilar consonants, presented visually or auditorily under concurrent articulatory suppression. All three give broadly the same results and will be reported in full subsequently as part of the wider study. For present purposes we have chosen one of the conditions in which participants experienced both visual and auditory presentation and performed suppression at a regular rhythm, as was the case for Jones et al. (2006; Jones et al., 2004).

EXPERIMENT 1

Method

Participants

A total of 24 undergraduate students participated to partially fulfil a requirement for their introductory psychology class.

Memory materials and presentation

Participants saw or heard six letters on each trial. On half the trials the letters were phonologically similar, *CDGPTV*, and on half the trials they were phonologically different, *BFHJQR*. The letters were presented in different orders on each trial, and the similar- and different-letter strings were presented in a quasi random order, with two lists of each type in every four trials. All participants saw and heard the letters and lists in the same order. In the visual condition letters appeared one at a time in the centre of a computer monitor at the rate of one per second. In the auditory condition, each letter was recorded in a female voice, and enough silence was added to each sound file so that the letters could be presented at the rate of one per second. A delay of 10 s followed the presentation before the word RECALL appeared on the computer screen to cue participants to begin recalling the order of the letters.

Procedure

Participants attended two experimental sessions, one week apart. Participants saw the letters during one session and heard the letters during the other session. The order of auditory and visual presentation was counterbalanced across participants. Each condition was preceded by 4 practice trials, followed by 24 test trials, consisting of 12 lists with each type of letter. During the presentation of the letters and the 10-s delay interval following presentation, participants spoke the word MA at the rate of about twice a second. After the 10-s delay, participants wrote responses on an answer sheet containing six lines for each trial and wrote something on each line before filling in the next line.

Results

The results are shown in Figure 2. The percentage of letters correct in each serial position was analysed with a 2 × 2 × 6 (presentation mode by phonological similarity by serial position) completely within-subjects analysis of variance (ANOVA). While there was a clear interaction between mode of presentation and phonological similarity, $F(1, 23) = 28.98$, $p < .001$, $\eta^2 = .558$, there was no three-way interaction among these three factors, $F(5, 115) = 1.55$, $p = .18$. A test of simple effects of phonological similarity at each serial position with each type of presentation showed that, with visual presentation, there was no phonological similarity effect at any location along the serial position curve, but with auditory presentation there was a significant phonological similarity effect at all serial positions except the first position.

Discussion

As Figure 2 shows, contrary to the findings reported by Jones et al. (2004), there is a clear phonological similarity effect, which is significant, when presentation is auditory at all serial positions except Position 1. While our results clearly differed from those of Jones et al. and appeared to provide clear support for the phonological loop model, an alternative explanation suggested itself. Overt articulatory suppression could have provided acoustic masking of the stimuli that could be particularly disruptive of perception of the similar letters. A study by Baddeley (1968, Exp. 1) suggested that this was unlikely to be a major effect, but given that the experimental materials and conditions used were very different from those in the present study, this clearly presented an interpretative problem.

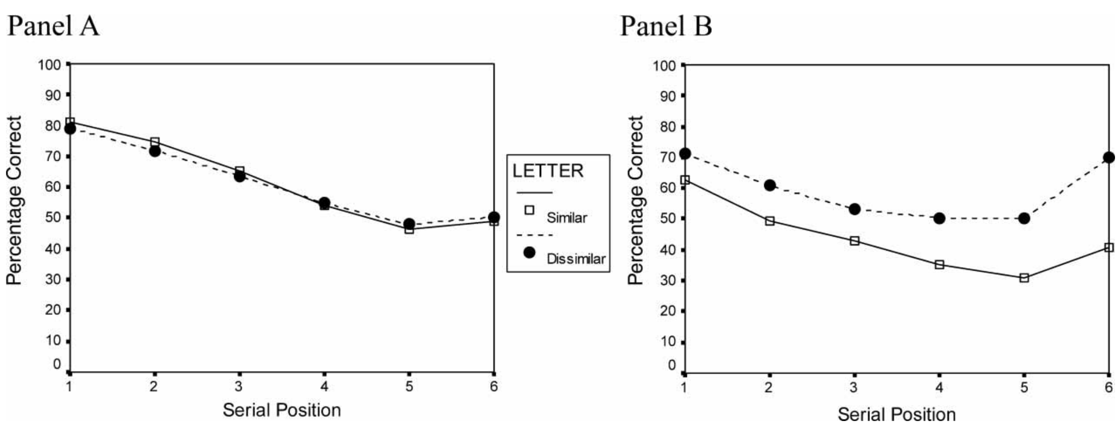

Figure 2. *Mean percentage of correct serial recall of similar and dissimilar letters with articulatory suppression. Panel A shows performance with visual presentation, and Panel B shows performance with auditory presentation.*

We therefore include a second study in which overt articulation is replaced by silent mouthing, which was shown by Murray (1967) to be broadly equivalent to overt suppression in its effects on recall.

EXPERIMENT 2

Method

Participants

A total of 24 people who had not taken part in the previous study participated to receive credit toward an introductory psychology requirement.

Memory materials and procedure

All features of this study were identical to those describe above, except that participants silently mouthed the word MA rather than speaking it aloud.

Results

The results are shown in Figure 3. The percentage of letters correct in each serial position was analysed with a $2 \times 2 \times 6$ (presentation mode by phonological similarity by serial position) completely within-subjects ANOVA. With silent mouthing, there was a clear interaction between mode of presentation and phonological similarity, $F(1, 23) = 25.312$, $p < .001$, $\eta^2 = .524$. In this experiment there was a three-way interaction among the factors, $F(5, 115) = 2.44$, $p = .04$. A test of simple effects of phonological similarity at each serial position with each type of presentation showed that, with visual presentation, there was no significant phonological similarity effect at any location along the serial position curve (Figure 3, Panel A), but with auditory presentation there was a significant phonological similarity effect at all serial positions (Figure 3, Panel B).

In this case, the interaction with serial position does support the assertion by Jones et al. (2006; Jones et al., 2004) that there is a contribution to performance from an auditory modality effect, which principally reflects the more recent items and is reduced by the presence of phonological similarity.

Discussion

Our results thus resemble those of Murray (1968) rather than the findings of Jones et al. (2006; Jones et al., 2004). All these studies agree that, with visual presentation, articulatory suppression disrupts the phonological similarity effect. A clear effect of similarity across serial positions returns when presentation is auditory in our data and in data of Murray, but not in data of Jones et al. Our silent mouthing results do resemble those of Jones et al. in showing a small recency effect with visual presentation and a larger one when presentation is auditory, with the effect being stronger for dissimilar than similar items.

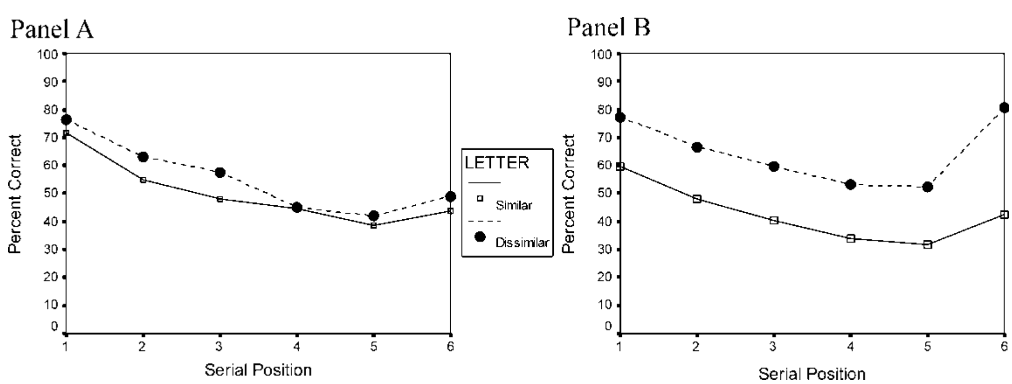

Figure 3. *Mean percentage of correct serial recall of similar and dissimilar letters with silent mouthing. Panel A shows performance with visual presentation, and Panel B with auditory presentation.*

Despite the 10-s gap, the standard auditory modality effect remains, as the literature suggests it should, while being reduced by the effects of phonological similarity as observed by Crowder (1971a, 1971b), Darwin and Baddeley (1974), and by Jones et al. (2006). In our own data, however, this is a very small contributor to a substantial effect. Given that the effect survives a 10-s delay filled by articulatory activity, it seems unlikely that it is either perceptual or "gestural" in nature.

One unexpected feature of our data is that performance is better with visual than with auditory presentation, whereas the typical trend is for the opposite. It is possible that the presence of articulatory suppression and phonological similarity may have encouraged participants to make greater use of supplementary, visual, or possibly semantic codes, which may be more readily accessible with visual presentation. Evidence for visual coding within the standard verbal recall paradigms is discussed by Logie, Della Sala, Wynn, and Baddeley (2000). A more worrying possibility is some form of acoustic masking of the stimulus sequences by the act of articulatory suppression. If this were the case, however, we would expect to have a different pattern of results when suppression involves silent mouthing. We found no such difference.

We suggest that our results do indeed limit the generality of the claim that the crucial interaction is limited to the recency effect and suggest that they do therefore compromise the proposal by Jones et al. (2006; Jones et al., 2004) to account for short-term memory entirely in terms of perceptual factors plus "gestural" articulation. We see other reasons for rejecting their view that information storage is unnecessary, but require more space to reply to the many pages of discussion offered by Jones et al. (2006; Jones et al., 2004) than is appropriate for a short empirical report. Our results do, however, require at least some consideration as to why the use of phonological coding declines as list length and level of difficulty increase.

One possibility is strategic. With visual presentation, converting to a phonological code is clearly optional, but presumably may have some attentional cost. As such a code becomes increasingly unproductive it may well be abandoned. In the case of auditory presentation, phonological coding is assumed to be automatic. However, retrieval on the basis of such a phonological code is presumably optional. When lists are long, and rehearsal is disrupted by suppression, then there may be little information to be gleaned from a decayed set of phonological traces. Note that this would still apply under conditions such as those used by Jones et al. (2006), in which the subject cannot predict in advance the nature of the next sequence. Given the theoretical importance of the disappearance of the phonological similarity effect as level of difficulty increases, it is clearly high time to study this relationship directly, using both strategy instructions, as did Hanley and Bakopoulou (2003), and ideally employing neuroimaging to check that subjects are indeed following the prescribed strategy.

A second issue of immediate relevance is the question of why the prefix and suffix conditions should, as we suggest, be regarded as effectively adding to list length, rather than influencing only perception as Jones et al. (2006; Jones et al., 2004) propose. Available evidence suggests that both prefixes and suffixes influence performance in two separate ways. One is by delaying recall. Conrad (1958) showed that prefixing recall of a number sequence with a key press was effectively like adding an additional item, while Baddeley and Hull (1979) showed that prefixing or following a digit sequence with a long or short Welsh town name (*Abergavenny* or *Rhyl*) impaired recall, with long names causing more disruption than short. Importantly, this effect had two contrasting components. Long suffixes impaired recall throughout the list, but had less impact on recency than did short suffixes. This result resembles that of Crowder (1978) who found less disruption from a repeated than a single suffix. We suggest that it is this latter recency-based effect that has formed the basis of the crucial interaction in the studies by Jones et al., while the former through-list storage component is the principal source of the interaction between suppression, modality, and similarity in our own findings.

Why should an auditory prefix that is not spoken by the subject influence performance? The phonological loop model assumes that spoken material gains automatic access to the store, potentially causing the prefix to be encoded as the first item of the sequence. Although redundant, it would still make demands on the limited encoding resources assumed by computational models of the phonological loop such as those proposed by Burgess and Hitch (1999) and Page and Norris (1998). The same could apply in the case of a suffix, particularly when these are spoken in the same voice as the items to be recalled as in Experiment 3 of Jones et al. (2006). The fact that current computational models do not yet deal specifically with such perceptual factors may be a limitation, but is surely not a reason for their rejection.

In conclusion, we suggest that Jones et al. (2006; Jones et al., 2004) have identified a genuine component of the Crucial Modality × Suppression × Similarity interaction in verbal short-term memory, but that it is a small effect relevant mainly to performance when the phonological loop is overloaded and/or abandoned. It forms a very limited basis for their rejection of the phonological loop hypothesis, which continues to receive extensive support from the study of performance under more standard conditions, and which has proved fruitful in throwing light on a wide range of issues from language acquisition (Baddeley, Gathercole, & Papagno, 1998), through neuropsychology (Vallar, 2006), to action control (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003) and education (Pickering, 2006).

Original manuscript received 2 March 2006
Accepted revision received 30 April 2006
First published online 12 February 2007

REFERENCES

Baddeley, A. D. (1966a). 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. (1966b). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology*, *18*, 302–309.

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

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

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

Baddeley, A. D., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, *130*, 641–657.

Baddeley, A. D., Gathercole, S. E., & 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. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D., & Hull, A. J. (1979). Prefix and suffix effects: Do they have a common basis? *Journal of Verbal Learning & Verbal Behavior*, *18*, 129–140.

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

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.

Conrad, R. (1958). Accuracy of recall using key set and telephone dial, and the effect of a prefix digit. *Journal of Applied Psychology*, *42*, 285–288.

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

Conrad, R., & Hull, A. J. (1964). Input modality and the serial position curve in short-term memory. *Psychonomic Science*, *10*, 135–136.

Crowder, R. G. (1971a). The sound of vowels and consonants in immediate memory. *Journal of Verbal Learning & Verbal Behavior*, *10*, 587–596.

Crowder, R. G. (1971b). Waiting for the stimulus suffix: Decay, delay, rhythm, and readout in immediate memory. *Quarterly Journal of Experimental Psychology*, *23*, 324–340.

Crowder, R. G. (1978). Mechanisms of auditory backward masking in the stimulus suffix effect. *Psychological Review*, *85*, 502–524.

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics*, *5*, 365–373.

Darwin, C. J., & Baddeley, A. D. (1974). Acoustic memory and the perception of speech. *Cognitive Psychology*, *6*, 41–60.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, *48*, 148–168.

Hall, J. W., Wilson, K. P., Humphreys, M. S., Tinzmann, M. B., & Bowyer, P. M. (1983). Phonemic similarity effects in good vs. poor readers. *Memory and Cognition*, *11*, 520–527.

Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review*, *10*, 435–444.

Hanley, J. R., & Broadbent, C. (1987). The effects of unattended speech on serial recall following auditory presentation. *British Journal of Psychology*, *78*, 287–297.

Hintzman, D. L. (1967). Articulatory coding in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, *6*, 312–316.

Johnston, R. S. (1982). Phonological coding in dyslexic readers. *British Journal of Psychology*, *73*, 455–460.

Jones, D. M., Beaman, P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 209–237). Hove, UK: Psychology Press.

Jones, D., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language*, *54*, 265–281.

Jones, D., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 656–674.

Larsen, J., & 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.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate serial recall. *Quarterly Journal of Experimental Psychology*, *53A*, 626–646.

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.

Neath, I., Bireta, T. J., & Surprenant, A. M. (2003). The time-based word length effect and stimulus set specificity. *Psychonomic Bulletin & Review*, *10*, 435–444.

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

Pickering, S. E. (Ed.). (2006). *Working memory and education*. New York: Academic Press.

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

Vallar, G. (2006). Memory systems: The case of phonological short-term memory. A festschrift for Cognitive Neuropsychology. *Cognitive Neuropsychology*, *23*, 135–155.

Wickelgren, W. A. (1969). Auditory or articulatory coding in verbal short-term memory. *Psychological Review*, *76*, 232–235.