

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/174037/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Roodenrys, Steven, Miller, Leonie M., Guitard, Dominic and Neath, Ian 2025. Similar phonemes create interference in the serial recall task. *Memory* 33 (2) , pp. 248-258. 10.1080/09658211.2024.2433049

Publishers page: <https://doi.org/10.1080/09658211.2024.2433049>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Similar Phonemes Create Interference in the Serial Recall Task

Steven Roodenrys¹

Leonie M. Miller¹

Dominic Guitard²

Ian Neath³

¹ School of Psychology, University of Wollongong, NSW, Australia

² School of Psychology, Cardiff University, Cardiff, United Kingdom.

³ Department of Psychology, Virginia Tech, Blacksburg, VA, USA

Author Note

The data and stimuli are available from the Open Science Foundation

(https://osf.io/ftgza/?view_only=8687f2c1e7f24cfe998a4fdb3f0e6d3f). Correspondence can

be sent to any author at steven_roodenrys@uow.edu.au, leoniem@uow.edu.au,

GuitardD@cardiff.ac.uk, ineath@vt.edu.

Word count: 8391

Abstract

The mechanisms underlying forgetting have been central to theorising about verbal short-term and working memory, and the importance of interference as opposed to decay continues to be vigorously debated. Here, we present two experiments to evaluate the nature and locus of phonological interference as a source of forgetting in serial recall. In these experiments, we replicate studies showing that repetition of phonemes across items impairs recall of the later list item, even with visual presentation and typed recall. In addition, we manipulate the degree of phonemic overlap between a target word and preceding words in the list, and demonstrate that consonants that are only similar, as opposed to identical, to each other can interfere with the recall of other list items. Taken together, the experiments suggests that the primary factor driving phonological interference is the similarity of the articulatory features of the phonemes.

KEYWORDS : short-term memory, serial recall, interference, output interference

The Nature of Phonological Interference in Serial Recall Tasks

The question of why people forget has been central to the study of memory from the beginning (see, e.g., Burnham, 1888, for an early review). Within the short-term/working memory literature, two causes have been frequently suggested (for a review, see Neath & Surprenant, 2003): Passive decay of information over time and active interference from other list items. The current paper seeks to further our understanding of the latter by further examining a particular type of interference that has recently been identified (Roodenrys, Miller, et al., 2022; James & Roodenrys, 2023).

Cowan et al. (2002) provided evidence that the strong primacy of the serial position curve (better recall of items near the beginning of the list) is partly due to output interference. They presented participants with strings of 9 digits to recall but varied the position in the list from which they were to start recall, and how many to recall. For the present purposes, the important conditions required participants to recall all 9 items but to start recall from position 1, 4 or 7, and for the latter two starting points to “wrap-around” and keep recalling in order from position 1 after position 9. The results are complicated because they also manipulated presentation modality and grouping of the stimuli into sets of 3, but, in general, there was a strong benefit to an item being recalled earlier during the output, irrespective of the input position.

Some models of short-term memory instantiate output interference by adding random noise to the system when an item is recalled. In the C-SOB model (Lewandowsky & Farrell, 2008) of serial recall and the SOB-CS model of complex span (Oberauer et al., 2012), when an item is output, Gaussian noise is added to the weights connecting a layer of units representing list-position markers and a layer representing the items. The interference

demonstrated by Cowan et al. (2002) was with the constrained set of digits, which might be thought to limit errors to the order of recall, however the SOB-CS model can simulate errors with more open sets of stimuli because in that model when an item is retrieved the noisy item vector is compared to a set of candidate items in long-term memory and a single item is selected. If the stimuli come from a well-learned, pre-defined set, such as digits, then only those items would be involved in this comparison, and if the stimuli come from an open pool of words it may be that a much larger set of representations in long-term memory are involved.

In contrast, it has also been shown that output interference can be specific, in a manner that cannot be modelled by adding random noise to processes or representations. Roodenrys, Miller et al. (2022) showed that having the three phonemes of a consonant-vowel-consonant (CVC) target word present in prior words in the list reduced the likelihood of recall. For example, a target word “bat” preceded by “bed, cap, nut” would be recalled more poorly than when it was preceded by words that did not share any phonemes with it. Experiments in this paper utilised the procedure of Cowan et al (2002) to dissociate order of recall and order of presentation by having participants recall some lists starting with position 4 and “wrapping around” to the start of the list to recall all 6 positions. It was argued that the combined results of the experiments strongly supported the view that the interference arose during recall rather than encoding.

James and Roodenrys (2023) replicated this effect when all three phonemes of the target occurred in prior list words, however, they also included conditions with only some phonemes of the target occurring earlier in the list. In one condition a target word was preceded in the list by words with the consonants of the target only, “b” and “t” in the previous example. In another condition only the vowel of the target occurred in another list word prior to the target. They found presenting both the consonants or only the vowel

created an equivalent degree of interference with the target, and slightly less than half the interference when all three phonemes were presented. In a further experiment, they demonstrated that prior presentation of the two consonants of the target only created interference if they occurred in the same intra-syllabic position as they did in the target word. That is “bed” and “nut” interfered with the recall of “bat”, but “hub” and “tin” did not. These findings demonstrate a specific output interference based on the similarity of the list items, in terms of the phonological elements of the words.

The Feature Model (Nairne, 1988, Nairne, 1990; Neath & Nairne, 1995) and Revised Feature Model (Saint-Aubin et al., 2021) includes a specific interference mechanism that operates during encoding, whereby if a word being encoded shares a feature with an earlier word in the list, that feature can be overwritten in the representation of the earlier word. Oberauer and Kliegl (2006) argued that the interference process could be competitive such that when two words shared a feature, either, but only one of them, could encode that feature as part of the representation of that word in the list. The data reported by Roodenrys, Miller et al. (2022) showed only very strong proactive interference effects within a list and are therefore inconsistent with specific interference amongst features working retroactively within the list or in a competitive manner.

The experiments reported by Roodenrys, Miller et al. (2022) and James and Roodenrys (2023) demonstrate that the recall of words containing the same phonemes as a target word awaiting output from short-term memory dramatically interferes with the recall of the target. Both of these studies used spoken recall by the participants. Unpublished studies conducted in the same lab found this interference effect also occurs with visual presentation and typed recall. One aim of the first experiment reported here is to demonstrate that the interference effect is not limited to spoken output. Such a finding would suggest that the effect does not

arise solely within the peripheral articulatory processes in recall but that processes more central to remembering the items are also responsible.

More importantly, this experiment also investigates whether an interference effect can be seen if phonemes that are only similar to those of the target, rather than the same as those of the target, are present in words prior to the target in the list. If this is found to be the case, it would suggest that the interference effect arises from the involvement in serial recall of representations coding information below the level of the phonemes. In order to examine this possibility, it is necessary to use a definition of the features of phonemes that allow for a clear identification of which phonemes are similar. For the purposes of this study, a definition based on articulatory features is adopted and applied only to the consonants of the words as this type of coding system provides especially clear examples of similar phonemes.

Experiment 1

The definition of phonemes in terms of a unique combination of articulatory features has a relatively long history (e.g. Chomsky & Hale, 1968; Ladefoged, 1971). These are often referred to as distinctive features and include such things as whether the vocal cords vibrate or not, the part of the mouth that is most open, and, if the vocal tract is fully closed during production, where that closure occurs. Using this system of classification, it is possible to identify phonemes that differ on a single feature. The term “minimal pair” is usually used to refer to two words that differ on a single phoneme, but it is used in this paper to refer to two phonemes that differ on a single feature.

Both consonants and vowels can be classified in this way but there is a smaller number of minimal pairs amongst the vowels, and they often differ in whether the vowel is tense or not, meaning they differ in length of pronunciation. A word length effect based on pronunciation duration, poorer recall for list of words that take longer to pronounce (Baddeley et al., 1975) has been a major feature of the short-term memory literature

contributing to theory development, but it has been strongly challenged by failures to replicate with different sets of words (e.g., Caplan et al., 1992; Guitard et al., 2018; Jalbert et al., 2011a,b; Service, 1998). James and Roodenrys (2023) demonstrated that interference was strongest when all three phonemes of the CVC target occurred in earlier words in the list but the interference on the target was also significant when just the vowel was shared with an earlier word in the list. Importantly for the present studies, the presence of only the two consonants of the target in earlier words also produced significant interference in recall of the target. It is possible to reduce any contentious possible confound with differences in duration in this experiment by manipulating only the overlap of the consonants of the targets with other words in the list and not the vowel. Although consonants also vary in length, unlike the vowels, a minimal pair manipulation of the consonants is not confounded with length of pronunciation.

Method

Participants. Sixty undergraduate students from the University of Wollongong Psychology participation pool participated in the experiment for course credit. They had a mean age of 20.3 years ($SD = 3.7$ years), and 49 were female. Participants completed the online experiment at a time of their own choosing and the experiment took approximately 10 to 15 minutes to complete. They read information about the experiment before participating and provided consent by continuing with the online task. This research was approved by the Social Sciences Human Research Ethics Committee of the University of Wollongong (Approval number 2022/080).

James and Roodenrys (2022) reported data for a condition where the target shared the consonants with two other words in the list and showed an effect of $d = 0.6$. A sensitivity analysis for a directional t-test was conducted with G*Power (Faul et al., 2009; Version: 3.1.9.6), and showed 60 participants provides a power of .92 to detect a more moderate effect

of $d = 0.4$. James and Roodenrys (2022) had 15 lists per condition and 32 participants in a within-subjects design. A larger number of participants in the present experiment is also warranted due to the smaller number of trials per condition.

Stimuli. All the stimuli used were CVC words (see appendix). There were 30 target words which were always presented in position 5 in the list, and which condition a target word was in was counterbalanced across participants such that each target occurred equally often in each condition. The same filler words were presented with a target, regardless of which condition the target was in, in positions 1, 3 and 6. The different interference conditions were created by presenting specific words in positions 2 and 4. In the overlap condition the words in positions 2 and 4 shared the initial and final consonant with the target respectively. For example, “ban” (bæn) and “wig” (wɪg) for the target “beg” (bɛg). In the minimal pair condition, the initial consonant of the word in position 2 and the final consonant of the word in position 4, differed from the consonants in the target by a single articulatory feature. For example, “pan” (pæn) and “wick” (wɪk) for the target “beg” (bɛg). For 36 of the 60 consonants of the targets, the minimal pair consonant differed on voicing, meaning that they are produced with the same mouth position/movements except that in one phoneme the vocal cords vibrate and the other they do not, e.g. the “b” and “p” or “k” and “g” (hard g) phonemes, as in the example above (see appendix for full list).

In the control condition the words in positions 2 and 4 did not share any phonemes with the target, and the consonants were not minimal pairs of the consonants in the target. In the control and minimal pair conditions, no phonemes were repeated within a list, while in the overlap condition only the overlapping phonemes of the target were repeated within the list. No word was presented to a participant more than once.

The words used in position 2 across the three conditions were matched on frequency according to the SUBTlex database (Brysbaert & New, 2009) and the number of letters (all *ps*

> .59 in a paired-samples t-test). The words used in position 4 were matched on frequency (all p s > .8) but unfortunately we were not able to match on the number of letters. The words in position 4 in the control condition were significantly longer (mean = 4.4 letters) than those in the overlap (mean = 4.0 letters) and minimal pair conditions (mean = 3.9 letters) (both p < .02) which did not differ from each other (p > .58). It should be noted, however, that if this difference of less than half an additional letter, on average, impairs recall of the following target word it will reduce the likelihood of finding significant interference effects because it will affect the control condition rather than the interference condition.

Procedure. A bespoke program for running serial recall over the internet was used to run the experiment. On each trial, participants saw a fixation cross at the centre of the screen and clicked a button on the screen to start the trial. They were then presented with six words in white on a black background, one after the other, in the centre of the screen at a rate of one word per second. Once the last word had been presented, it was replaced on the screen by a response box labelled “stimulus 1” and participants typed their response before clicking another button to submit the response and move on to the next word. They were instructed that they could leave a response box blank when they could not recall the corresponding word. Thirty-two trials were presented in a single block. The first two trials were practice trials, and the remainder were the 30 experimental trials, presented in a different random order for each participant.

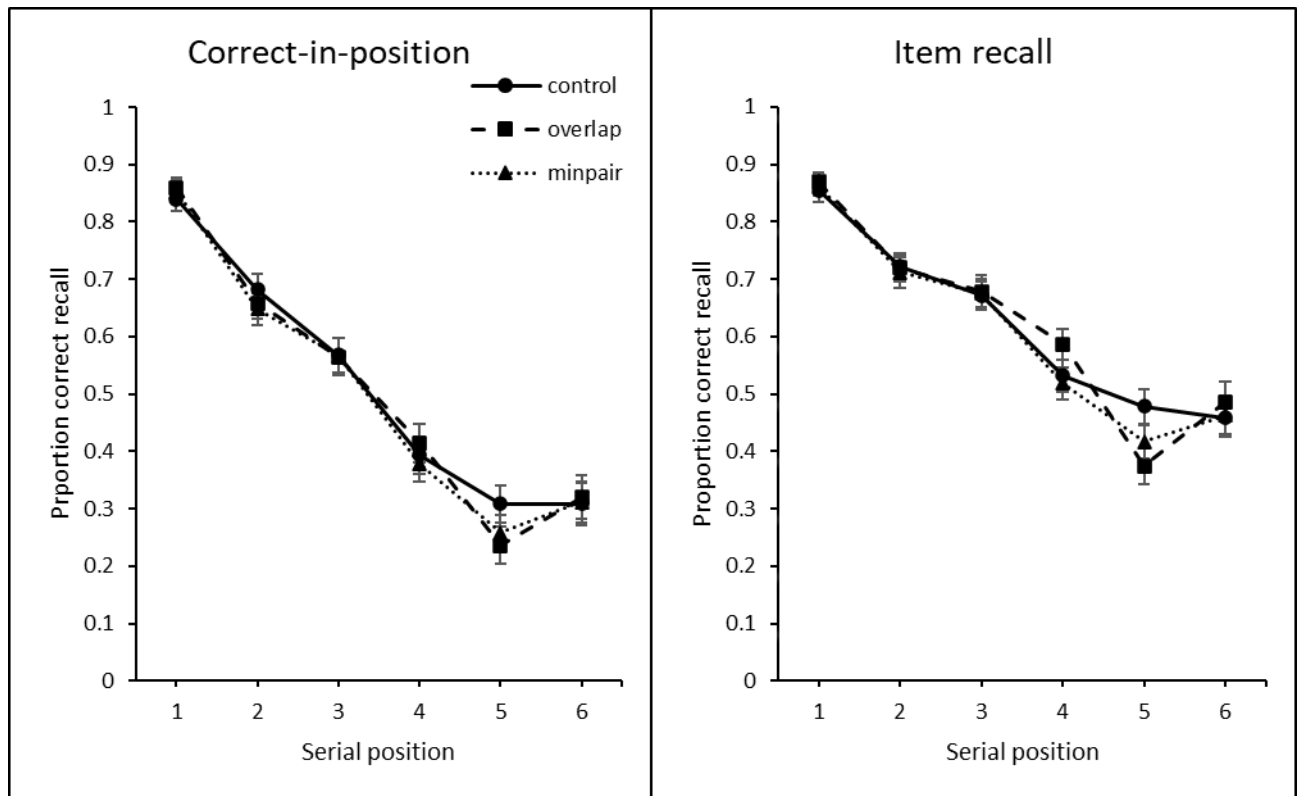
Results

Prior to analysis of responses, typographical and spelling errors were corrected. These were homophones, adjacent letter transpositions, hitting an adjacent key on the keyboard resulting in a nonword being entered, or a spelling error that preserved the phonology. Approximately 1.2% of responses were corrected but only 0.07% were target words (one

target in the control condition, three in the overlap condition and two in the minimal pair condition). Two measures of recall were calculated, a correct-in-position measure, which required the word to be recalled in the position in which it was presented, and an item-recall measure, which allowed recall of the word in any position in the list. The means for these measures across each serial position are shown in Figure 1. These measures were subjected to both frequentist and Bayesian statistical analyses using JASP (JASP team, 2023), as described below.

Figure 1

Serial recall data from Experiment 1



Note. The left-hand panel shows the data scored using a correct-in-position criterion, while the right-hand panel shows the item-recall data, i.e. scored without respect to position. The error bars represent SEM.

To evaluate whether recall prior to the target items was equivalent, a (4) x (3) repeated-measures ANOVA on the correct-in-position data from the first four serial positions and the three conditions was conducted. This analysis showed a significant effect of serial position, $F(3,177) = 178.624$, $MSE = 0.036$, $p < .001$, $\eta_p^2 = 0.752$, but no effect of condition, $F(2,118) = 0.340$, $MSE = 0.031$, $\eta_p^2 = 0.006$, or condition by position interaction, $F(6,354) = 0.821$, $MSE = 0.014$, $\eta_p^2 = 0.014$. Compared to a model that included the serial position factor, there was very strong evidence for the null hypothesis ($BF_{10} = 0.05$) when the model also included

condition, and extremely strong evidence for the null ($BF_{10} = 0.0005$) when the model included both condition and the condition by position interaction. As can be seen in the lefthand panel of Figure 1, these analyses show recall was equivalent across the three conditions in the first four serial positions.

Directional paired t-tests were used to evaluate the predictions that recall of the target would be worse in the two interference conditions than the control condition, and worse in the overlap than the minimal pair condition. In the correct-in-position data, the first test showed that the target was significantly less well recalled in the overlap condition ($M = .237$) than in the control condition ($M = .308$), $t(59) = 3.125$, $p < .01$, Cohen's $d = .403$, $BF_{10} = 21.629$, providing strong evidence for an interference effect. Recall was also significantly worse in the minimal pair condition ($M = .258$) than the control, $t(59) = 2.01$, $p < .05$, Cohen's $d = .260$, $BF_{10} = 1.792$, providing anecdotal level evidence for an interference effect. The difference between the minimal pair and overlap conditions was not significant, $t(59) = 1.089$, Cohen's $d = .14$, $BF_{10} = 0.424$, providing moderate level evidence in favour of the null effect.

The item-recall data for the targets were also subjected to directional t-tests. The target was significantly less well recalled in the overlap condition ($M = .374$) than in the control condition ($M = .478$), $t(59) = 3.502$, $p < .01$, Cohen's $d = .452$, $BF_{10} = 59.732$, providing very strong evidence for an interference effect. Recall was also significantly worse in the minimal pair condition ($M = .417$) than the control, $t(59) = 2.338$, $p < .05$, Cohen's $d = .302$, $BF_{10} = 3.428$, providing moderate level evidence for an interference effect. The difference between the minimal pair and overlap conditions failed to reach significance, $t(59) = 1.507$, Cohen's $d = .195$, $BF_{10} = 0.763$, providing anecdotal level evidence in favour of the null effect.

Discussion

The results provide a clear replication, with a new set of stimuli, of James and Roodenrys' (2023) demonstration that the prior inclusion of just the consonants of the CVC target in other words in the list is sufficient to produce a notable level of interference. Whereas the previous studies used auditory presentation and spoken recall, this experiment demonstrates the interference effect is observed even when presentation is visual as opposed to auditory and responses are typed as opposed to spoken, suggesting the effect is not tied to processes involved in the overt spoken production of the words.

The results also provide some evidence that this proactive interference effect of phonemes within the list is not restricted to identical phonemes. In the minimal pair condition, where very similar consonants to those in the target occurred in words earlier in the list, recall of the target was significantly impaired, with anecdotal level evidence for the correct-in-position scoring and moderate level evidence for the item-recall scoring. Although this evidence is not strong enough to draw firm conclusions at this point, this result has implications for the way phonemes are represented within the systems underlying serial recall. Before discussing the results in greater detail, a second experiment is presented.

Experiment 2

The second experiment is a replication of the first, using the same stimuli, with the exception that they were presented auditorily and recall was spoken. This experiment was conducted to see if these changes in the procedure would provide stronger evidence for an interference effect in the minimal pair condition, where the consonants are similar to each other rather than identical. We reasoned that spoken recall in particular, should increase the

strength of output interference as it maximises the involvement of articulatory phonology, even if participants are using other strategies to try and recall the lists.

Method

Participants. Sixty different undergraduate students from the University of Wollongong Psychology participation pool participated in the experiment for course credit. They had a mean age of 20.4 years ($SD = 5.0$ years), and 50 were female, and 1 non-binary. They provided written consent before completing the experiment.

Stimuli. The stimuli were the same as in Experiment 1. The words were recorded by a male Australian English speaker using a Røde® Broadcaster studio condenser microphone and Audacity® 2.1.1 software. Each stimulus was normalised to control for possible amplitude effects on recall and saved as a separate file for use with the experimental control software. SuperLab Pro V5.0 was used to present the lists via an external amplified speaker.

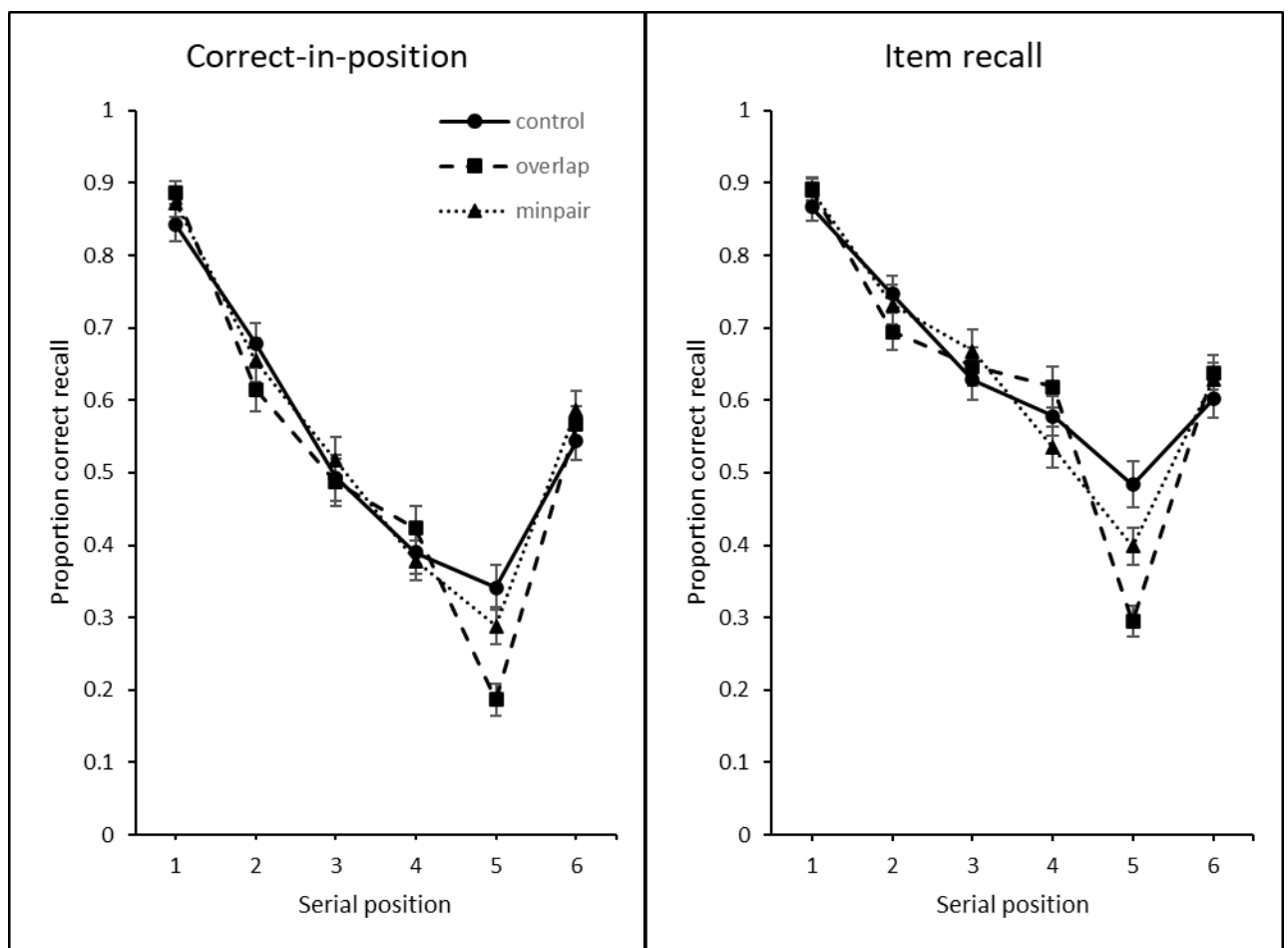
Procedure. Testing was conducted in a small laboratory room with the experimenter controlling the presentation of the lists and recording responses using a prepared response sheet. On each trial, participants heard the six words of the list presented at a rate of 1 word per second. They were instructed to begin recalling the list in the order of presentation beginning with the first word, once the list had been presented. They were told they could say “blank” for any words they could not remember. The first two trials were practice trials, after which the experimenter checked with the participant that they were comfortable with the task before proceeding to the experimental trials. The 30 experimental trials, which were identical to Experiment 1, were then presented in a different random order for each participant.

Results

The data were analysed in the same manner as in Experiment 1. The means for each condition by serial position are presented in Figure 2.

Figure 2

Serial recall data from Experiment 2



Note. The left-hand panel shows the data scored using a correct-in-position criterion, while the right-hand panel shows the item-recall data, i.e. scored without respect to position. The error bars represent SEM.

As in Experiment 1, a (4) x (3) repeated-measures ANOVA on the correct-in-position data from the first four serial positions and the three conditions was conducted to determine if recall was equivalent across the conditions prior to the target word. This analysis showed a significant effect of serial position, $F(3,177) = 252.093$, $MSE = 0.030$, $p < .001$, $\eta_p^2 = 0.612$, but no effect of condition, $F(2,118) = 0.350$, $MSE = 0.036$, $\eta_p^2 = 0.000$. There was a significant condition by position interaction, $F(6,354) = 3.716$, $MSE = 0.013$, $p = .001$, $\eta_p^2 = 0.0048$. Compared to a model that included the serial position factor, there was strong evidence for the null hypothesis ($BF_{10} = 0.042$) when the model also included condition, but only anecdotal evidence for the null ($BF_{10} = 0.0426$) when the model included both condition and the condition by position interaction. As can be seen in the lefthand panel of Figure 2, these analyses show recall was equivalent across the three conditions in the first four serial positions, although the curve for the overlap condition shows a slightly more scalloped shape than the other conditions.

In the correct-in-position data, a directional paired t-test showed that the target was significantly less well recalled in the overlap condition ($M = .187$) than in the control condition ($M = .342$), $t(59) = 6.711$, $p < .001$, Cohen's $d = .866$, $BF_{10} = 2.915 \times 10^6$, providing extreme evidence for an interference effect. Recall was also significantly worse in the minimal pair condition ($M = .288$) than the control, $t(59) = 2.413$, $p < .01$, Cohen's $d = .312$, $p < .05$, $BF_{10} = 4.013$, providing moderate level evidence for an interference effect. The difference between the minimal pair and overlap conditions was also significant, $t(59) = 4.653$, $p < .001$, Cohen's $d = .601$, $BF_{10} = 2022.371$, providing extreme evidence of a stronger interference effect in the overlap condition.

The item-recall data for the targets was evaluated using the same directional t-tests. The target was significantly less well recalled in the overlap condition ($M = .295$) than in the

control condition ($M = .483$), $t(59) = 6.489$, $p < .001$, Cohen's $d = .838$, $BF_{10} = 1.288 \times 10^6$, providing extreme evidence for an interference effect. Recall was also significantly worse in the minimal pair condition ($M = .398$) than the control, $t(59) = 3.035$, $p < .01$, Cohen's $d = .392$, $BF_{10} = 17.184$, providing strong evidence for an interference effect. The difference between the minimal pair and overlap conditions was also significant, $t(59) = 3.843$, $p < .001$, Cohen's $d = .496$, $BF_{10} = 159.436$, providing extreme evidence of greater interference in the overlap condition.

Discussion

The pattern of the results in Experiment 2 replicates those of Experiment 1 very closely. In line with the aim of the experiment, the interference effects are stronger in Experiment 2 than in Experiment 1. This can be seen clearly in the figures, where there is much greater separation between the conditions in the recall of the target, and also in the effect sizes. For example, the difference in recall of the control and minimal pair conditions in the item recall data showed $d = .302$ in Experiment 1 and $d = .392$ in Experiment 2. The results of this experiment provide more convincing evidence that the presence of phonemes that are only similar to – as opposed to identical to – those of the target word earlier in the list can interfere with recall of the target word.

General Discussion

Previous studies (Roodenrys, Miller et al., 2022; James & Roodenrys, 2023) have demonstrated that the same phoneme occurring in the same intra-syllabic position across different words in a serial recall task produces interference during recall of the list. The experiments reported here provide the first demonstration that interference between specific phonemes in different words in short-term memory occurs when the phonemes are simply

similar to each other rather than identical. This phenomenon offers a novel constraint on theories, and our understanding of what drives forgetting in short-term memory.

These results provide evidence for interference in short-term memory when a phoneme is repeated across items within a list using a new set of stimuli generated under different constraints from the previous studies. As such, it increases confidence that specific interference occurs during output of the lists rather than merely the addition of random noise to the system, as some models do in order to simulate human performance (e.g. Oberauer et al., 2012). Roodenrys, Miller et al. (2022) suggested that this interference may be the result of suppression of recalled items spreading through a multi-level interactive activation network to the phonemes that they are comprised of. Response suppression has been widely adopted in models that involve competition between items to be output, in order to prevent the model from repeatedly recalling the same item (e.g. Page & Norris, 1998). The results reported here are consistent with an account in which suppression at recall of an item affects more than just the activation of the representation at the item level, the item vector in many models. If suppression is applied to the item representation, in a multi-level, interactive model it also results in the inhibition of the representations of the phonemes that make up the item. If the different layers have bi-directional links. The graded effects reported in these experiments, where phonemes that are similar to those in the target produce less interference than the identical phoneme, suggests that the articulatory features of phonemes are also represented in the system. Suppression of the features will then reduce the activation of the phonemes of the target, although not by as much as when the target phonemes themselves have been presented. The result is a reduction in the activation of the representation of the target item when it comes to be recalled, reducing the likelihood of correct recall.

However, Kowialiewski and Oberauer (2024) have recently suggested the results of Roodenrys, Millet et al. (2022), and by the implication the results reported here, could be

explained by overwriting at recall. That is, when an item is recalled, the elements of the representation of that item that match those in a later list-item awaiting recall may be deleted from the representation of the later item. Apart from proposing this as an alternative to suppression in order to explain the interference effects seen in these experiments, they do not offer an explanation for why such a mechanism might exist. If the two representations that share elements have co-existed in memory up until the point of recall, how would the overall operation of the memory system be improved by overwriting of the representation remaining in memory? As Kowialiewski and Oberauer (2024) acknowledge, response suppression was suggested as a mechanism to explain memory phenomena such as the relative infrequency of repetition errors in serial recall. They report the results of experiments in which participants were asked to recall items from a list multiple times, and argued that if response suppression exists, the likelihood of the second recall should be reduced compared to the first. They found equivalent performance on the two recall occasions and therefore reject the existence of response suppression. They do suggest some alternatives to explain their results involving the additional use of episodic memory, but these are not related to the interference effects in the present paper. In Kowialiewski and Oberauer's (2024) experiments participants were cued to recall the list items by position, but in random order. It may be that this task is sufficiently different from serial recall to involve different output processes, as participants know they will not need to output the list in order even if they have to remember the positions of the items. Further work is clearly required to investigate this. Based on the available evidence, we suggest there is a stronger rationale for the existence of response suppression in serial recall, spreading through the articulatory features that make up the phonemes, than overwriting at recall.

The novel finding of these experiments is that interference between phonemes in different words can occur between phonemes that are very similar rather than identical.

Similarity was defined in terms of the distinctive articulatory features (Chomsky & Hale, 1968) of the phonemes rather than acoustic similarity. Phonemes that have similar articulatory features can be acoustically similar but this is not always true. Only about half of the minimal pair phonemes in the stimuli set are acoustically confusable (Wang & Bilger, 1973). However, if the effect was due to acoustic confusability it would be reasonable to expect it to be absent from Experiment 1 which did not use auditory presentation of the stimuli. The extremely similar results of the two experiments argues against the effect being driven by acoustic similarity. In comparison, it is not surprising that articulatory processes are active in Experiment 1. There has been extensive research on the nature of phonological representations activated when reading silently which suggests they are at least partly articulatory (see Leinenger, 2014, for a review) and articulatory processes have been shown to support typing (Hayes & Chenoweth, 2006). Of course, articulatory processes have long been seen as playing a major role in the serial recall task, even when presentation is visual. Repeating an irrelevant phrase during list presentation (articulatory suppression - e.g. Levy, 1971) reduces the accuracy of recall, and the relevance of articulatory rehearsal has long been accepted (Baddeley & Hitch, 1974). However, as Experiment 1 was conducted online, it is possible that participants read the words aloud, resulting in acoustic processing of the stimuli, and further work to more clearly rule out the possibility of the interference being acoustically based is required. More specifically, and following Roodenrys, Miller et al. (2022), articulatory suppression could be utilised to minimise both overt articulation of items during presentation and the possibility of interference arising from rehearsing items prior to output. Roodenrys, Miller et al. showed that engagement with articulatory suppression did not alter interference effects from the repetition of phonemes across items, and, by extension, preservation of the interference effects seen with the minimal pair manipulation would be expected.

It should also be acknowledged that repetition of phonemes is confounded with repetition of letters in Experiment 1, and even the minimal pair condition involves the presentation of similar letters in the interfering words to those of the target (e.g. “p” and “b”). It may be possible to construct lists that repeat phonemes without repeating letters, as English includes phonemes that can be represented by two different letters (e.g. a “soft g” makes the same sound as “j”, a “hard c” makes the same sound as “k”) but this would greatly limit the number of potential stimuli.

The reader can examine the visual similarity between the minimal pair stimuli in the Appendix. In order to see if visual similarity might be driving the interference effect in the minimal pair condition, we combined the two letter pairs for each target to form three groups (both similar, both dissimilar, and one dissimilar and one similar). The stimuli pairs mostly involve one similar and one dissimilar letter, making it impossible to run statistical comparisons, and a visual examination of the data does not suggest a clear pattern. However, the consistency of the results between Experiment 1 with visual presentation and typed recall, and Experiment 2 with auditory presentation and spoken recall could suggest that visual similarity does not play a strong role in the effect. Nevertheless, future research might test the possibility that all or part of the effect is driven by visual similarity by using articulatory suppression throughout the recall period.

The interference effects observed were larger in Experiment 2 than Experiment 1, in line with the aim behind changing the procedure. The intention was to see if stronger evidence for interference in the minimal pair condition than was provided by Experiment 1 could be obtained by maximising the role of articulatory phonology in the task. The studies were not designed to examine a modality effect, so the magnitude of the interference in the two experiments has not been directly compared. This is a possible direction for future research.

The interference seen in these experiments from having similar consonants to those of the target word earlier in the list suggests that the effect arises from the use of articulatory representations of the items during recall. James and Roodenrys (2023) showed that interference between phonemes across words occurred only if they occupied the same position within their CVC word stimuli. They suggested that the use of a syllable template to represent each word, as in the Hartley and Houghton (1996) psycholinguistic model, could explain this intra-syllabic position-specific interference between words. The results of these experiments suggest that phonemes in such a model may be represented as a set of articulatory features, and suppression of a phoneme may result in the suppression of other phonemes with a very similar set of features.

Interactive activation models of a number of language processes exist, so the suggestion that the activation of articulatory features as part of such a system plays a role in recall has some precedent. Although the phonological interference effect in the serial recall task is argued to arise from articulatory processes, there is evidence in the speech perception literature that the recent processing of a phoneme hinders the identification of a similar phoneme. Newman and Dell (1978) showed that when participants monitored sentences for the occurrence of a target consonant phoneme, the detection of that phoneme was slower if the preceding word contained a similar phoneme. In their experiment, the target was always the “b” phoneme at the start of a word, and the preceding word either started with the “p” phoneme (which is a minimal pair phoneme for it), another stop consonant, or a non-stop consonant, which are the most different from “p”. Detection was slowest following the minimal pair phoneme, quicker following the other stop consonants, and fastest following the non-stop consonants. They argued that this provides evidence that articulatory features play a role in the perception of phonemes. Stemberger et al. (1985) ruled out the possibility of misidentification of the first phoneme delaying identification of the target phoneme, and

instead explained the interference effect in terms of an interactive activation model. They suggested that identification of the target phoneme was delayed when a similar phoneme had been presented shortly beforehand because the similar phoneme was still active. The similar phoneme would receive some activation from the shared features when the target was presented, allowing it to compete for longer than if it had not been presented. It would be possible to argue for a similar effect in the minimal pair condition in the experiments reported here, with the recently produced phoneme competing with the phoneme of the target word resulting in poorer recall. However, competition from a recently activated phoneme cannot explain the interference effect of a repeated phoneme in the overlap condition, where one would expect the correct phoneme to be even more activated. It therefore seems that response suppression, spreading at the level of articulatory features, offers a more parsimonious explanation for the pattern of results in the serial recall task.

The interference effects investigated here, along with some other work described below, have implications for one of the major findings in the short-term memory literature, the phonological similarity effect (PSE). The PSE is the finding that lists of words which are all similar to each other are recalled less well than lists which are dissimilar (Wickelgren, 1965). A common claim has been that this is due to a disturbance of memory for the order of the items but not the identity of the items. Papers using different operational definitions of similarity, such as all the words in a list rhyming, versus a high degree of phonological overlap (e.g. *mad, man, map, mat, cad, can, cap, cat*), have supported this claim, although the evidence is more complicated when open pools of stimuli have been used instead of small, closed sets of stimuli (see Roodenrys, Guitard et al, 2022, for a review). The lack of a PSE on memory for the items in a list seems contrary to the findings of the experiments reported here, however Roodenrys, Guitard et al. (2022) recently provided data addressing this issue. They compared recall of lists of rhyming words with lists in which two phonemes of each

word occurred in another word in the list, but there were no phonemes that were shared by all items in the list (e.g. *lime, heat, line, wing, hoot, ding*). They found that item recall was poorer in the latter condition compared to a dissimilar condition in which none of the words shared any phonemes. They argued that lists where all the words share some component such as the rhyme or the vowel, provide a cue to the identity of the items. Different components of the word may vary in their effectiveness as a cue, and along with other factors may explain variation in the degree to which they have been able to overcome the effect of interference in the studies in the literature.

Regardless of whether these results reflect a suppression mechanism or the overwriting of features of items in memory, the evidence that overlap in articulatory features is sufficient to result in interference between words in serial recall raises the possibility that this is the mechanism for all output interference. These experiments created a specific relationship between the features of a pair of phonemes earlier in the list and two phonemes in the target word, but in any list used in a serial recall task there will be some degree of overlap in articulatory features between words. The further a word is positioned into a list, the more often the features of its phonemes are likely to have occurred in previous items. Models such as C-SOB (Lewandowsky & Farrell, 2008) produce a primacy gradient in serial recall through the accumulation of random noise added to the representations of the items remaining to be recalled. The results of these experiments suggest that the degradation arising during output may not be random, but a function of overlap in the articulatory features of the items. The lists used in these experiments were not constructed to control for the amount of overlap in features across all items in the list and the target word. They manipulated the overlap between specific phonemes and ensured no additional repetition of phonemes within a list. Future work may be able to test the possibility that output

interference accumulates based on the number of times features have already been part of the earlier words in the list.

In conclusion, similarity at the level of articulatory features of items has been shown to produce interference-based forgetting in serial recall. The presence of interference is not dependent on processes specifically related to spoken output and is argued to likely reflect the operation of response suppression during the recall of items. As such, given the likelihood that featural overlap exists across items in lists used widely in the literature, it raises the possibility that loss of memory otherwise assumed to be a function of random noise in the system could in fact, be a principled form of forgetting.

Declaration of interest

The authors report no conflict of interest.

References

- Baddeley, A., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). Academic Press.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, 14(6), 575-589. doi:10.1016/S0022-5371(75)80045-4
- Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990. doi:10.3758/BRM.41.4.977
- Burnham, W. H. (1888). Memory, historically and experimentally considered. I. An historical sketch of the older conceptions of memory. *American Journal of Psychology*, 2(1), 39-90. Doi:10.2307/1411406
- Caplan, D., Rochon, E., & Waters, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 45A(2), 177–192. doi:10.1080/14640749208401323
- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.
- Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding serial recall. *Journal of Memory and Language*, 46(1), 153-177. doi: 10.1006/jmla.2001.2805

- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. doi:10.3758/BRM.41.4.1149.
- Guitard, D., Saint-Aubin, J., Tehan, G., & Tolan, A. (2018). Does neighborhood size really cause the word length effect? *Memory & Cognition*, 46(2), 244–260. doi:10.3758/s13421-017-0761-9
- Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*, 35(1), 1-31. doi:10.1006/jmla.1996.0001
- Hayes, J. R., & Chenoweth, N. A. (2006). Is Working Memory Involved in the Transcribing and Editing of Texts? *Written Communication*, 23(2), 135-149. doi:10.1177/0741088306286283
- Jalbert, A., Neath, I., Bireta, T. J., & Surprenant, A. M. (2011a). When does length cause the word length effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 338-353. doi:10.1037/a0021804
- Jalbert, A., Neath, I., & Surprenant, A. M. (2011b). Does length or neighborhood size cause the word length effect? *Memory & Cognition*, 39(7), 1198-1210. doi:10.3758/s13421-011-0094-z
- James, T., & Roodenrys, S. (2023) Exploring the necessary conditions for phonological interference in serial recall. *Memory*, 31(7), 891-904. doi: 10.1080/09658211.2023.2200587
- JASP Team (2023). JASP (Version 0.18.1.0) [Computer software] <https://jasp-stats.org/>

- Kowialiewski, B., & Oberauer, K. (2024). Testing the response suppression mechanism of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi:10.1037/xlm0001359
- Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. University of Chicago Press.
- Leinenger M. (2014). Phonological coding during reading. *Psychological Bulletin*, 140(6), 1534-55. doi:10.1037/a0037830
- Levy, B. A. (1971). Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning & Verbal Behavior*, 10(2), 123–132. doi:10.1016/S0022-5371(71)80003-8
- Lewandowsky, S., & Farrell, S. (2008b). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 49, pp. 1-48). Elsevier Academic Press.
- Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition*, 16(4), 343-52. doi: 10.3758/bf03197045
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, 18(3), 251-269. doi:10.3758/BF03213879
- Neath, I., & Nairne, J. S. (1995). Word-Length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review*, 2(4), 429-441. doi:10.3758/BF03210981
- Neath, I., & Surprenant, A. M. (2003). *Human memory: An introduction to research, data, and theory, second edition*. Wadsworth.

- Newman, J. E., & Dell, G. S. (1978). The phonological nature of phoneme monitoring: A critique of some ambiguity studies. *Journal of Verbal Learning & Verbal Behavior*, 17(3), 359–374. doi:10.1016/S0022-5371(78)90228-1
- Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55(4), 601-626. doi:10.1016/j.jml.2006.08.009
- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, 19(5), 779-819. doi: 10.3758/s13423-012-0272-4
- Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3737–3753. doi:10.1098/rstb.2009.0173
- Roodenrys, S., Guitard, D., Miller, L. M., Saint-Aubin, J. & Barron, J. M. (2022). Phonological similarity in the serial recall task hinders item recall, not just order. *British Journal of Psychology*, 113(4), 1100-1120. doi: 10.1111/bjop.12575
- Roodenrys, S., Miller, L. M., & Josifovski, N. (2022). Phonemic interference in short-term memory contributes to forgetting but it is not due to overwriting. *Journal of Memory and Language*, 122, Article 104301. doi:10.1016/j.jml.2021.104301
- Saint-Aubin, J., Yearsley, J. M., Poirier, M., Cyr, V., & Guitard, D. (2021). A model of the production effect over the short-term: The cost of relative distinctiveness. *Journal of Memory and Language*, 118, Article 104219. doi:10.1016/j.jml.2021.104219

- 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 A: Human Experimental Psychology*, 51A(2), 283–304. doi:10.1080/027249898391639
- Stemberger, J. P., Elman, J. L., & Haden, P. (1985). Interference between phonemes during monitoring: Evidence for an interactive activation model of speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 11(4), 475–489. doi:10.1037/0096-1523.11.4.475
- Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. *The Journal of the Acoustical Society of America*, 54(5), 1248-66. doi: 10.1121/1.1914417. PMID: 4765809.
- Wickelgren, W. A. (1965). Short-term memory for phonemically similar lists. *American Journal of Psychology*, 78(4), 567-574. <https://doi.org/10.2307/1420917>

Appendix

Table A1: Stimuli lists used in both experiments.

Filler	C1 overlap	Filler	C2 overlap	Target word	Filler	C1 minpair	C2 minpair	C1 control	C2 control
Shed	Rice	Mug	Cap	Rip	Faith	Lice	Cab	Vice	Con
Head	Seek	Fool	Gap	Sip	Torn	Thatch	robe	Murk	Chug
Marsh	Lane	Porch	Bog	Lag	Third	Rough	Seek	Tone	Whizz
Sack	Ball	Verge	Heat	Bet	Watch	Dim	Heap	Loan	Harsh
Pun	Gem	Size	Vague	Jig	Shot	Charm	Hack	Leech	Mode
Vein	Botch	Surf	Luck	Bark	Shoot	Poach	Wig	Match	Loathe
Fizz	Dale	Pot	Germ	Dime	Youth	Bowl	Cane	Luck	House
Face	Bomb	Cheat	Wig	Beg	Van	Pool	Wick	Mole	Ward
Walk	Tail	Chef	Boom	Team	Sang	Pill	Gown	Gill	Buzz
Shop	Gill	Dice	Hem	Game	Woof	Cat	Horn	Hat	Bowl
Job	Thin	Sad	Calm	Thumb	Whirl	Tip	Hen	Vet	Pouch
Jazz	Bush	Void	Knock	Bike	Torch	Push	League	Hoop	Thin
Shack	Doubt	Worth	Halve	Dove	Gym	Base	Roof	Yell	Port
Zoom	Jug	Whack	Hop	Jeep	Soil	Chin	Height	Fog	Dine
Hug	Chain	Verb	Fort	Chart	Leash	Jam	Ripe	Fame	Yawn
Safe	Tongue	June	Lark	Tick	Burp	Thud	Log	Rash	Vile
Shape	Root	Mass	Weed	Rod	Zing	Lawn	Vote	Cuff	Hive
Vine	Jog	Shell	Curb	Jab	Race	Chick	Weep	Turf	Dock
Word	Reap	Sign	Buck	Rack	Jet	Loaf	Bug	Theme	Butch
Hedge	Calf	Peach	Board	Cod	Wish	Gate	Sub	Mile	Verse
Boon	Thief	Gel	Pig	Thug	Chat	Soap	Dock	Posh	Weave
Fuzz	Tame	Witch	Vogue	Tag	Herb	Doll	Soak	Veal	Sheen
Hung	Fed	Charge	Beat	Fate	Norm	Veal	Pad	Sob	Cash
Wheel	Sob	Shine	Math	South	Ridge	Zap	Coat	Beep	Curve
Bid	Mate	Gull	Nerve	Move	Pose	Neat	Cough	Fork	Heath
Wool	Tack	Rogue	Barn	Ton	Haze	Dab	Farm	Mash	Course
Mall	Pave	Yarn	Ruck	Poke	Food	Bathe	Rig	Sash	Hutch
Voice	Rage	Feed	Warp	Rope	Gosh	Latch	Web	Hawk	Gin
Less	March	Hook	Pod	Mud	Gash	Knife	Pit	Phone	Whip
Fetch	Balm	Thing	Wheat	Boat	Hall	Pack	Guide	Jive	Wash

The first six columns show the lists presented in the overlap condition. To make the minimal pair condition lists the words in the columns labelled “C1 minpair” and “C2 minpair” were substituted into positions 2 and 4. The same was done with the words in the “C1 control” and “C2 control” columns to create the control lists.

Table A2: Minimal pair phonemes used in the experiments.

Phoneme pair	Number of times used in stimuli	Type of feature difference
l/r	5	place
θ/s	2	manner
θ/t	3	manner
d/b	4	place
d/t	6	voice
tʃ/dʒ	4	voice
p/b	10	voice
p/t	3	place
k/g	12	voice
v/f	3	voice
z/s	1	voice
n/m	7	place