Testing feature interaction : Between stream irrelevant speech effects in immediate recall

Tolan, Georgina Anne and Tehan, Gerald

Tolan, G. A. and Tehan, G. (2002). Testing feature interaction : Between stream irrelevant speech effects in immediate recall. *Journal of Memory and Language*, 46, pp. 562-585. https://doi.org/10.1006/jmla.2001.2820

Testing Feature Interaction: Between-stream Irrelevant Speech Effects in Immediate Recall

Georgina Anne Tolan

Australian Catholic University

Gerald Tehan

University of Southern Queensland

Address correspondence to

Gerry Tehan
Department of Psychology
University of Southern Queensland
Toowoomba, QLD. Australia. 4350

e-mail: tehan@usq.edu.au

Abstract

Immediate serial recall of visually presented lists is disrupted by irrelevant background speech. One explanation for the irrelevant speech effect assumes that features of the auditory material become incorporated into the memory trace of the to-be-remembered item thereby reducing the fidelity of the short-term trace. From this perspective the resultant short-term memory trace is a composite of features of the list item and features of an item in the irrelevant stream. While there is evidence that item interactions in short-term memory are observable, there is currently no direct empirical evidence for such interactions involving irrelevant speech. We report six experiments using a short-term cued recall task that manipulates proactive interference in which item interactions have been observed. In these experiments we consistently show that with irrelevant speech specific items in the auditory stream influence target recall and the presence or absence of proactive interference. These results provide relatively direct evidence for feature interaction involving irrelevant speech. The results are evaluated against the three current models of irrelevant speech effects

Testing Feature Interaction: Between-stream Irrelevant Speech Effects in Immediate Recall

Ordered recall of short lists of items is disrupted by someone speaking in the background, even though the instructions stress that this speech is to be ignored. The irrelevant speech effect, as this phenomenon is known, has become an important benchmark in the development of theories of short-term recall.

Currently, there are three different explanations for the irrelevant speech effect. The working memory explanation (Baddeley, 1986) argues that the words in the unattended stream are registered in the same short-term store as the list items and this produces recall problems. At this point in time Baddeley has not specified what causes the difficulties. The initial versions suggested that irrelevant speech corrupted the phonological store in some non-specific way but this view was rejected in favour of an assertion that irrelevant items add some form noise to the recall process (Baddeley & Hitch, 1994). In the most recently published discussions of the phonological loop model, Baddeley still gives no clear indication of precisely how irrelevant speech produces a recall deficit (Baddeley, 2000).

The O-OER model (Jones, 1993) argues that the interfering effects of background sound are not limited to speech. Jones and his colleagues have consistently shown that tones produce the same pattern of disruption as speech. In both instances disruption is a function of the degree to which items in the irrelevant stream change from item to item; changing state stimuli produce more interference than steady state stimuli. In the O-OER model, the deficit from irrelevant sound is due to confusions in pathways that maintain the order of items in memory. When a list is learned the items are linked to each other to form a pathway from the first to the last item. In the case of irrelevant speech, the items in the auditory stream form a second pathway that links the irrelevant items together. At recall the participant attempts to recover the list items by following the pathway for those items. However, confusions in the linkages are possible which produce recall errors. Such cross talk between the two pathways is more likely with changing state sound than steady state sound.

In the Jones model the memory trace for the list items is not degraded in any way by the irrelevant speech. In the Baddeley model it is not clear how or if trace degradation occurs. In the third account, irrelevant speech directly affects the memory trace. Neath's (1999; 2000) adaptation of the feature model (Nairne, 1990) argues that features from the irrelevant speech become absorbed into the memory trace of the list items. That is, features of the words in the auditory stream interact with features of the list items such that the memory trace for the list item is changed. The resultant short-term trace loses its fidelity and exists as a corrupted version of the original trace. The corrupted short-term trace is then matched to a series of long-term traces to select an appropriate response. To the extent that the short-term trace is corrupted, the probability of selecting the correct item decreases in that there may not be a substantial match with any list item, or the match might be closer to an incorrect item than to the correct item.

In discussing Neath's approach, both Baddeley (2000) and Jones and Tremblay (2000) implicitly argue that there is no direct evidence for feature adoption. Furthermore, they suggest that at face value, feature adoption predicts that the similarity between the auditory items and the to-be-remembered items should have an impact upon performance. Both expect that the greater the level of similarity, the greater the resulting disruption. The available empirical evidence indicates that between-stream similarity has little impact upon recall. There are now close to a dozen separate experiments that have explore between-stream similarity effects (Bridges & Jones, 1996; Jones & Macken, 1995; Larsen, Baddeley & Andrade, 2000; LeCompte & Shaibe, 1997) and none, bar Salamé & Baddeley's (1982) Experiment 5, have found any evidence in favour of between-stream similarity. The implication is that this lack of evidence is problematic for the feature model. Although Neath is able to formally model the absence of between-stream similarity effects, neither Baddeley nor Jones and Tremblay are overly receptive to the assumptions Neath has made in order to do so. As such, the absence of between-stream similarity effects has the potential to be problematic for feature adoption. This issue might be partially resolved if between-stream similarity effects could be observed.

In the current research we test the ideas behind the Neath perspective, rather than the way feature adoption is formally implemented in the feature model. That is, the primary interest is in the possibility that items in memory may interact with each other and that a resultant memory trace may in fact be a blend of two or more items. We think that Baddeley and Jones and Tremblay are incorrect in assuming that feature adoption can only result in a detrimental effect upon performance. It may do so, but it is possible that it may have no effect or it is possible that irrelevant speech may facilitate recall under some circumstances. Some examples might make this clear.

The feature model involves a matching process in which an easily corruptible short-term trace is matched to a set of traces in long-term memory where both short-term and long-term traces consist of sets of features. The item that produces the best match is selected for recall. Let us take the simplest case where two words have been studied, say dog and cat, and the person is matching the short-term trace of cat to the long-term traces. If one were to do this task solely on the basis of the match between the letters of the respective words, it is readily apparent that there is complete overlap between the letters involved in the short-term trace (*c*, *a* and *t*) and the long term trace for cat (*c*, *a* and *t*) and no overlap with the long term trace for dog (*d*, *o* and *g*). Consequently, there should be little problem in recalling cat.

In the feature model forgetting is based upon retroactive interference where features in the short-term trace are overwritten. Now let us assume that the short-term trace has been overwritten such that only the middle letter is available (*a*). It is still the case that matching *a* to the features of *c*, *a* and *t* and *d*, *o* and *g* will still produce a better match for cat than dog so cat is again likely to be recalled.

Now assume that with unattended speech features of the auditory words are absorbed into the short-term trace. The simplest form that this could take is if some of the features in the irrelevant words replace the features of the short-term trace in their respective positions within a word (Neath does not make this assumption). Consider the effects of having *fib, cot* or *dig* in the auditory stream, a corrupted short-term trace in which only *a* is available, and the two consonants of the auditory items are incorporated into the short-term trace in their respective positions. In the case where *fib* is paired with the short-term trace, the resulting short-term trace will be *fab*. This trace will match with the long-term representation of cat on one of the three features and will not match any features of dog. Cat is likely to be recalled. In the case where *cot* is paired with the short-term trace, the *c* and the *t* will be absorbed such that the resultant trace is *cat*. This will produce a perfect match with the long-term trace of cat. Here we have an instance of where irrelevant speech restores the missing features and actually facilitates recall. In the case of *dig* in the irrelevant stream, the *d* and the *g* will be absorbed into short-term trace to form *dag*. Note now that the short-term trace produces two matches with dog and only one with cat; dog is more likely to be recalled. In this instance the irrelevant speech is likely to produce a recall error.

While the above examples maybe overly contrived, the general point is still valid. As Baddeley and Jones and Tremblay note, if items are represented in terms of features and features from different items are able to interact, then short-term recall should be sensitive to the type of features that are present in both memory and the irrelevant stream. However, it is not necessarily the case that the effects should be disruptive. It should depend upon the features involved. Sometimes list recall might be enhanced and in other instances it might be hurt by irrelevant speech. We now turn to evidence for the type of feature interaction effects that we are concerned with.

Feature (Item) Interaction Effects

Over recent years we have been exploring proactive interference (PI) effects in a short-term cued recall task (Tehan & Humphreys, 1995, 1996, 1998; Tolan & Tehan, 1999). In this task participants are presented a series of trials consisting of either one or two four-word blocks. Participants are instructed to pay attention to all the items, but once they find out that a trial is two-block trial, they are directed to forget the first block and concentrate on remembering the second block because it will be this block that will be tested. At test a category cue is presented and participants are asked to recall the item from the most recent block that was an instance of the category. The critical trials in the experiment are the two-block trials because it is on these trials that PI is manipulated. On the two-block trials there is always a target item presented amongst three unrelated filler items in the second block. In the control conditions, the target item is the only instance of the category in the list. In the interference condition, a second instance of the category is presented in the first block. Thus, the cue might be *ANIMAL,* the block-2 target might be *cat*, and the block-1 foil might be *dog*.

Using this paradigm we have consistently observed PI effects on a delayed test (Tehan & Humphreys, 1995, 1996; Tolan & Tehan, 1999). Our exploration of item interaction effects has centred on immediate recall where sometimes recall is immune to the effects of PI and in some instances PI can readily be observed. It turns out that the presence or absence of PI critically depended upon the phonological similarity between the foil and the target (Tehan & Humphreys, 1995, 1998). Recall was immune to PI when taxonomic categories were used and the foil and target did not rhyme. However, when word ending cues were used (*_AT* as the cue, *cat* as the target, and *hat* as the foil) or when rhyming instances of taxonomic categories were used (*ANIMAL, cat* as the target, and *rat* as the foil) PI effects were observed. Target recall was depressed and recall of the foil increased substantially.

Tehan and Humphreys (1998) tested for feature interaction by exploring the possibility that with non-rhyming instances of a taxonomic category, the features of the foil could be provided by another word. In these experiments they utilised standard control and interference trials but added a series of trials in which one of the filler items in the second block was a rhyme of the foil (*ANIMAL, cat* as the target, and *dog* as the foil, and *log* as a block-2 filler). When the rhyme was absent from the list, no effects of PI were observed. However, when the rhyme was in the list, PI effects were observed. The presence of *log* in the second block enhanced the recall of *dog*. They took the idea of feature interaction a step further by taking the phonemes of the foil and distributing these items across three filler items in the second block. Thus *dog* was decomposed into *__d__art, m__o__p* and *fi__g__*. This manipulation again produced an increase in PI, primarily through enhancing recall of the foil.

The strengthening of the foil by items elsewhere in the list that share features with it is what we mean by an item interaction. The above research suggests that at the point of recall the features representing the target, the foil and the filler items appear to be simultaneously active in memory and the phonemic features of these items appear to interact with each other. We think that results of this type provide reasonably direct evidence for feature interaction.

If features from filler items in a list can influence recall it is not a giant leap to suppose that irrelevant auditory items could have similar effects. That is, if *dart, mop* and *fig* were items included in the auditory stream and the *d, o* and *g* were absorbed into the memory trace, recall of the foil *dog* should be enhanced in the same way that it was when these words were filler items. In short, feature interaction assumptions predict that phonological similarity between the items in auditory and visual streams could enhance PI in the cued recall task in the same way and for the same reasons that irrelevant filler items in the list enhanced PI. As we indicated previously, the available empirical evidence indicates that between-stream similarity effects are not observable, at least as far as immediate serial recall goes.

The failure to replicate the between-stream similarity effect appears to present a major obstacle for feature interaction assumptions. It looks as though similarity of features has no effect whatsoever. However, to this point, irrelevant speech effects have not been studied using a cued recall task, nor has the impact of irrelevant speech on short-term PI effects been studied. The experiments we report here first look at the effects of irrelevant speech on the cued recall task, and then explore the effects of between-stream similarity on the cued recall task. To preview our results we show that between-stream similarity **does have** an impact upon the cued recall task in that PI effects are influenced by the irrelevant items in the auditory stream.

Experiment 1

This experiment explores the effects of irrelevant speech in a short-term cued recall task. In this experiment we wish to determine whether or not irrelevant auditory input has a detrimental impact upon cued recall. Irrelevant speech effects are readily observed in immediate serial recall where multiple responses are required and order must be maintained. In fact, Jones (1993) has argued strongly that irrelevant speech effects are limited to those tasks that involve maintenance of order. Since, there is no a priori reason for students to maintain order in the cued recall task, irrelevant speech effects may well differ to those in serial recall. Furthermore, irrelevant speech effects are less robust in tasks where a single response is required (Beaman & Jones, 1997).

Provided that irrelevant speech does have an impact the question remains as to whether or not changing-state speech produces more of a decrement than steady-state speech (Jones & Macken, 1995). Thus, in the current experiment we compare a quiet control to both steady-state and changing-state speech.

Finally, the disruptive effect of irrelevant speech seems to occur at storage or retrieval rather than at encoding (Baddeley & Salamé, 1986; Colle & Welsh, 1976; Colle, 1980; Miles, Jones & Madden, 1991). In the current experiment we present the irrelevant speech simultaneously with the list items and test immediately, or we present the visual items in the absence of speech but delay testing until two seconds of irrelevant speech has been presented after the final item in the list. Given prior findings, irrelevant speech effects should be equivalent in the two conditions.

Method

Participants

Forty introductory level psychology students participated in this experiment for course credit. The first twenty participants were assigned to the irrelevant speech during input condition while the remaining twenty participants performed under speech during rehearsal period delayed condition.

Materials

The first step in constructing the stimulus set was the creation of two mutually exclusive pools of words to serve as materials for the critical two-block trials in the experiment. One pool contained the items that would serve as targets on the critical cued recall trials, while the other acted as a source for filler items. Following the Tehan and Humphreys (1995; 1996, 1998) procedure, the target items were created by selecting one instance from thirty-six different taxonomic categories was from the South Florida Category Norms (McEvoy & Nelson, 1982). These items were low dominant items within the category's hierarchy. The filler word-pool consisted of 252 words that came from different taxonomic categories to those used in the selection of the critical trials. The remaining categories from the South Florida Category Norms (McEvoy & Nelson, 1982) and the Shapiro and Palermo (1970) category norms served as the sources of these items, with multiple items from each category being selected. This ensured that there was no overlap in category membership between critical and filler items, but it was possible that two items from a category could appear as filler items on a trial. Items for each trial were chosen without replacement from their respective stimulus pool.

Each subject studied a unique set of forty-eight trials. Each trial consisted or one or two four-word blocks with the 36 two-block trials being the critical trials in the experiment. To create each two-block trial, seven items were randomly selected without replacement from the filler pool. Four of these items were randomly assigned to the first block and the remaining three were assigned to the second block. A target item was then randomly sampled without replacement from the target pool. On half the trials this item was placed as the second item in the second block and on the other half it appeared in position three. The thirty-six two-block trials where then randomly assigned to the three irrelevant speech conditions such that equal numbers of trials were assigned to each condition.

The study also contained 12 one-block trials. The items for these trials were also selected from the Sth Florida norms but were selected such that there was no overlap in category membership with any of the items in the two-block trials. The target items in these lists tended to be placed in the first and last serial position and were always tested in the absence of irrelevant speech. The order of the 48 trials was randomised for each subject.

We generated the items for the irrelevant stream by creating 72 phonotactically legal non-words and then assigning three non-words to each of the 24 lists that were to be studied under irrelevant speech conditions. These non-words were either one or two syllables long. Each three non-word combination in the changing-state condition was phonemically dissimilar from each other (e.g., NUG-PROG-BULA) while the three non-word combination in steady-state condition was phonemically similar to each other (e.g., GEN-VEN-NEN). None of the non-words rhymed with any of the target items.

It should be stressed that our manipulation of steady-state speech is not truly steady-state; the first phoneme changes across words. True steady-state speech would involve a simple repetition of a single item. We have adopted the current type of speech because this was the version of low-level changing-state speech that Jones and Macken (1995) used when exploring between-stream similarity effects.

Procedure

At the beginning of the experiment subjects were informed that they would be studying a series of one-block and two-block trials in which a block consisted of four words. However, it was also stressed that at any one point in time they only had to remember that most recent block of four items. Consequently, if the trial was a two-block trial, and this was signified by the presence of an exclamation mark (!) as the block separator, they were to forget the first block and concentrate on remembering the second block because it would be on this block that they would be tested. They were told that one-block and two-block trials would be randomly interspersed throughout the experiment and that since they would not know in advance what type of trial it was, it was in their best interests to treat each trial as a one-block trial until they learned otherwise. They were also told that on some trials they would hear the experimenter repeating a series of words and that they were to do their best to ignore this material.

The events that the subjects were concerned with all happened in the bottom left hand corner of a computer monitor. Each trial began with a READY sign displayed for two seconds. The study items were then displayed individually in lower case at a rate of one word per second, and subjects were instructed to remain silent and not to move their lips throughout the presentation of the study items. On two-block trials, the block separator, (!), was presented for one second after the fourth word in the first block and before the first word in the second block. At recall, a category cue was presented in upper case for two seconds. On an immediate test the cue appeared immediately after the fourth item in the block. On trials in which a rehearsal period was employed, the screen went blank for two seconds after the final word had been presented and subjects were instructed to rehearse the last block of items. After 2 seconds the cue appeared in the same corner of the screen. With the appearance of the cue, participants were requested to verbally recall the item from the most recent block that was an instance of the category. Subjects had five seconds to make a response before the next trial began. The experimenter recorded the subject's responses (correct recall, intrusion errors, omissions, etc) on a hard copy of the subject's input file.

The experimenter provided the irrelevant speech in the experiment. On the trials where irrelevant speech was manipulated, the items were presented in the top right hand corner of the screen and the experimenter repeated the items for the time they remained on the screen. With items appearing simultaneously in the diagonal corners of the screen, there is potential for visual activity to be a source of distraction and a possible confound. However, with the seating arrangements used, the perceptual experience for both the subject and the experimenter was at most a flicker in the peripheral visual field. It was not possible to read the words, nor was it possible to identify that the visual material was in fact words. Participants reported that this irrelevant visual material did not distract them and many in fact did not notice it. They mentioned that they were too busy concentrating on the task at hand and ignoring what was being articulated in the background. Consequently, any effects of the nature of the words in the irrelevant stream are unlikely to be due to the visual events that were occurring in the top part of the screen.

In the case of irrelevant speech that was presented during input, the three non-words appeared in the top right hand corner as the first to-be-remembered word appeared in the bottom left hand corner and stayed on the screen until the final list word disappeared. During this time the experimenter repeated the three non-words at a rate that produced about eight or nine repetitions of the non-word sequence during the 9 seconds it took to present the study list. When the irrelevant speech was presented during the two-second rehearsal period, the three non-words appeared in the top corner immediately after the final to-be-remembered disappeared from the bottom corner. The non-words were present for two seconds and the experimenter pronounced the sequence only once. Thus, the amount of irrelevant speech differed in the two conditions. The non-word sequence was repeated eight or nine times in the immediate test condition but only once in the rehearsal period condition.

Results and Discussion

One-Block Trials

The aim of these trials was to ensure that participants attend to the first block of words on all trials. The participants correctly recalled 86% of block-one trials in this experiment and performed at similar levels in the remaining experiments. As such these trials appeared to achieve their purpose and are not reported in subsequent experiments.

Two-Block Trials

Figure 1 displays the mean proportion of correct recall and errors made while attempting target recall under quiet, changing-state and steady-state conditions during input and rehearsal phases. However, in the interests of space analyses are only conducted on target recall. In places during the remainder of the manuscript we will make passing reference to the number of omission errors made. Where the text indicates that there are differences in omissions in the various conditions, these differences are statistically reliable. An alpha level of .05 has been used in all statistical analyses.

Correct recall A 2 x 3 repeated measures analysis of variance was conducted on the target recall. A main effect for when irrelevant speech was presented was evident, F (1, 38) = 15.09, MSe = 0.05. A significant main effect for irrelevant speech also emerged, F (1, 38) = 61.06, MSe = .01 Planned comparisons indicated that correct recall was significantly superior under the quiet condition than the steady-state, t (39) = 4.77, and the changing-state conditions, t (39) = 6.77. However, the changing-state and steady-state condition did not differ from one another, t (39) = .08. A time of presentation by type of irrelevant speech interaction emerged, F (1, 38) = 14.78, MSe = 0.01. It is clear from Figure 1 the irrelevant speech effects were stronger in the rehearsal phase condition than in the input phase condition. Importantly, however, simple effects analyses indicated that irrelevant speech effects were evident in both conditions. Thus, in the input phase performance under the quiet condition was significantly greater than under the changing-state, t (19) = 3.59, and the steady-state condition, t (19) = 2.36. Correct recall for steady-state and changing-state irrelevant speech did not differ from one another under the input conditions t (19) = 1.00.

For the rehearsal period performance was significantly superior under the quiet condition than steady-state, t (19) = 5.01, and changing-state conditions, t (19) = 7.02. Changing-state and steady-state conditions did not significantly differ from one another, t (19) = .69.

The results are quite straightforward. Irrelevant speech has a detrimental impact upon cued recall just as it does on immediate serial recall and a number of other memory tasks. Secondly, the effects of irrelevant speech are more pronounced when the auditory input occurs during a rehearsal period rather than during presentation. This result is somewhat surprising given that more irrelevant speech is experienced when the auditory material is presented during study than when it is presented in a rehearsal period. It might be the case that even though subjects are requested to rehearse during the rehearsal period, some forgetting is taking place. For instance on the quiet condition, recall after a two second unfilled retention interval is not as good as on an immediate test. Furthermore, with irrelevant speech it is possible that the auditory material effectively makes this rehearsal period a filled retention interval. Tolan and Tehan (1999) have shown that very brief periods of verbal distractor activity produce a marked negative impact upon the cued recall task.

Alternatively, from an O-OER perspective, irrelevant speech effects are expected to be produced where subjects are relying upon order information. In the delayed test students are instructed to rehearse the item whereas no such instructions were given for the immediate test group. To the extent that subjects are more reliant upon rehearsal in the delayed condition than in the immediate condition, irrelevant speech effects should be more prominent after a delay.

Finally, and perhaps most importantly, both steady-state and changing-state irrelevant speech appear to produce equivalent levels of disruption and this is true irrespective of when the irrelevant speech is presented. This finding is atypical in that changing-state speech usually produces a robust decrement, but steady-state speech produces either a weak effect or no effect at all. Since the current results are somewhat unexpected we thought it best to replicate the findings before placing too much weight on the current results. Thus in the next experiment we again manipulate the type of irrelevant speech we employ. The effects of irrelevant speech on proactive interference are also examined.

Experiment 2

Experiment 1 clearly demonstrated the effects of irrelevant speech under cued recall procedures. From a feature interaction perspective the irrelevant speech has its effect by becoming incorporated into memory. Here we assume that the primary effect of the irrelevant speech is on the phonemic representations in memory. Thus, the phonemic features of list items will be embedded within a background that contains the phonemic features of the irrelevant items. Tehan and Humphreys (1995, 1996; 1998) have argued that the phonemic representations of list items serve as a means of producing immunity to PI. Thus, if these representations are degraded by the irrelevant speech, their ability to assist in isolating the most recent memories from earlier memories may well be reduced. Thus, in the current experiment we revert to the PI version of the cued recall task. On some trials an interfering foil is presented in the first block. We expect that under quiet conditions, the presence of the foil will have little impact because the phonemic features of the target are retrieved against a background of minimal trace degradation and these features will reinforce the target. That is, we expect to observe immunity to PI. However, if irrelevant speech has an impact upon the phonemic features of the target, it is possible that target recall will become more difficult and that discriminating between the target and the foil will be more difficult. PI should be observed under irrelevant speech conditions.

We again use both steady-state and changing-state irrelevant speech, although in this experiment it is manipulated between subjects. Furthermore, the irrelevant speech was only presented during a two-second rehearsal period. Experiment 1 indicated that placement of the irrelevant speech during the rehearsal phase was a more potent manipulation even though the patterns performance were identical for both methods of presenting the auditory material.

Methods

Participants

A total of sixty participants took part in this experiment and were randomly assigned to one of three group according to order of arrival (n = 20 per group). The groups differed on the type of irrelevant speech they experienced.

Methods

The selection of materials and list construction was similar to that used in Experiment 1. This time, two instances from forty taxonomic categories from the South Florida norms were selected. Again following Tehan and Humphreys (1995) the target item tended to be a low dominant item from with in the category and the second which served as the foil in the interference trials, was a relatively high dominant instance from within the category, although never the most dominant. The filler items were selected from the same sources and care was taken so that there was no category overlap between the fillers and the target/foils.

Fifty-two lists, of which forty were two-block trials and twelve were one-block trials were constructed in the same way as in Experiment 1, save that on half the two-block trials, one of the filler items in the first block was replaced by the foil. The foil and the target were always allocated to the same serial position in their respective blocks; half the time in the second position and half the time in position three. Again each subject was given a unique set of trials in which the materials had been randomly allocated to the different trials and the order of the trials had been randomised as well.

The irrelevant speech items were the same as those used in Experiment 1.

Procedure

The procedure was identical to Experiment 1 except that irrelevant speech was manipulated between subjects. In the changing-state and steady-state conditions two groups of twenty participants were each exposed to irrelevant speech during a two-second rehearsal phase that appeared immediately after the fourth TBR item in the second bock of words and before presentation of the category cue. For the quiet condition the irrelevant speech was absent during the two-second rehearsal phase. It was, however, emphasised in all conditions that the two-second retention interval was to be used to rehearse the four TBR items and that the concurrent irrelevant speech should be ignored.

Results

Figure 2. presents the data for target recall and the errors made. PI effects can be determined either by differences in target recall or by the change in block-1 intrusions. Both recall measures were analysed.

Correct recall The correct recall data was submitted to a 2 x 3 mixed design analysis of variance, with interference being a within-subject manipulation and irrelevant speech a between-subjects factor. A significant main effect for irrelevant speech was obtained, F (2, 57) = 5.90, MSe = 0.02. Planned comparisons revealed that recall of the target item was significantly better under quiet conditions than steady-state, t (38) = 2.33, and changing-state, t (38) = 3.57, conditions. Target recall for changing-state and steady-state conditions did not significantly differ from one another, t (38) = 1.03. The main effect for interference was not significant, F (1, 57) = .57, MSe = 0.01, confirming that there was no difference in correct recall between interference and non-interference conditions. In addition, there was no significant interaction between irrelevant speech and interference, F (2,57) = 1.51, MSe = 0.01.

Block-one intrusions Planned comparisons indicated that in that significantly more foils were produced under the changing-state condition compared with the quiet condition, t (38) = 2.57. The steady-state condition also produced significantly more block-one intrusions than the quiet condition, t (38) = 2.19. The steady-state and changing-state conditions did not significantly differ from one another in the amount of block-one intrusions produced, t (38) = .14.

Discussion

Recall performance for the target item showed an irrelevant speech effect. Furthermore, steady-state and changing-state speech appeared to produce roughly equivalent levels of interference. There were some signs that changing-state speech resulted in more omission errors than steady-state speech, but in Experiment 1 this was not the case. Likewise, there were no reliable differences in block-1 intrusions between the two forms of irrelevant speech. Thus, the results of Experiment 2 replicate those of Experiment 1 by showing that there is an irrelevant speech effect in the cued recall task and that steady-state and changing-state speech produce equivalent levels of disruption.

The above description of the differences between steady-state and changing-state effects is based on instance by instance comparisons. It is possible to do an experiment wide meta-analysis of the results as well. We first calculated the effect size for the difference between the quiet and the steady-state means for all conditions in both experiments. Effect sizes based on the mean differences between quite and changing-states were also calculated and all effects sizes are reported in Table 1. Inspection of Table 1 suggests that the effects of changing-state speech are stronger than the effects of steady-state speech in all conditions. This difference was statistically reliable, t (8) = 7.26, p < .000. Hence this further analysis of the combined data supports the notion that changing-state irrelevant material does have a greater disruptive effect on memory performance than steady-state irrelevant material, although obviously the differential effects are much weaker in the cued recall task than they are in immediate serial recall. We explore this issue further in the General Discussion.

With regards to PI effects in the experiment, there was no significant difference between interference and non-interference conditions for the correct recall data. But this appears to be the result of different error patterns. That is, while the total number of errors are roughly equivalent in the three conditions, the frequency of the different types of errors change. Thus, more omission errors are made in the no-interference conditions than the interference conditions and this is offset by increased block-1 errors in the interference conditions. Given that block-one intrusions are the strongest measure of PI effects, the data indicate that irrelevant speech results in recall being vulnerable to PI. The speech interferes with the processes that produce immunity to PI in the cued recall task. There is substantial evidence that irrelevant speech eliminates the phonological similarity effect (Colle & Welsh, 1976; Jones & Macken, 1995; Salamé & Baddeley, 1982). This suggests that irrelevant speech interferes with phonological codes to the extent that the available phonological information is no longer of a sufficient quality that it unambiguously supports the target item.

Experiment 3

Experiments 1 and 2 confirm that short-term cued recall is affected by irrelevant speech, that both steady-state and changing-state produce an irrelevant speech effect, and that irrelevant speech results in increased levels of PI. In Experiments 3 to 6 the principal aim of the paper, that between-stream similarity has a detrimental impact upon cued recall performance, is tested. In the remaining experiments, save for Experiment 3, two changes in methodology are introduced. Firstly, items in the irrelevant stream now rhyme with the target or they rhyme with the foil. The similarity manipulation is crossed with an interference manipulation to produce a two by two factorial design. Thus, the irrelevant speech is either similar or dissimilar to a critical item and the critical item is either the target or the foil. This means that the irrelevant speech rhymes with an item on the list on three quarters of the trials: interference and non-interference trials where the speech rhymes with the target and interference trials in which the speech rhymes with the foil. However, on non-interference trials the irrelevant speech rhymes with a potential foil, even though that foil does not appear in the list.

Straightforward analyses of the factorial design are complicated if a quiet condition is added. Thus the second methodological change is to discontinue using a quiet control condition. The results of Experiments 1 and 2 clearly demonstrate that an irrelevant speech effect is readily obtained in the cued recall task. Our aim in the following experiments is solely with the issue of whether or not between-stream similarity influences recall performance. Whether or not performance is above or below a quiet control is irrelevant to the hypotheses under consideration.

From a feature interaction approach the predictions are reasonably simple. By having irrelevant speech sharing phonemic features with a list item, the irrelevant speech should replace some degraded features of the list item. To some extent this should restore the representation of that item and thereby increase the likelihood that it will be recalled. Thus, if the auditory material rhymes with the target, it should restore any missing features of the target item. In so doing, it should further protect the target from possible interfering effects of the foil. However, if the speech rhymes with the foil, the foil's representation should be restored or strengthened. This should lead to a decrease in target recall and an increase in block-1 intrusions. Thus, if the feature interaction assumptions are correct, there should be differences in target recall on both interference and non-interference trials. Secondly, block-1 intrusions should be more prevalent when the speech reinforces the foil than when it reinforces the target. However, if between-stream similarity is not an important component of the irrelevant speech effect, as much of the data would indicate (Bridges & Jones, 1996; Jones & Macken, 1995; Larsen et al., 2000; LeCompte & Shaibe, 1997), then there is no reason to expect that there would be differential effects on target recall or block-1 intrusions. It would still be expected that PI effects would emerge due to the irrelevant speech, but they should not differ as a function of the nature of the material in the auditory stream.

The use of irrelevant speech that rhymes with a list item introduces a potential problem. If the participants notice the relationship between the irrelevant speech and the task material, they could use that knowledge to predict the target item. A simple strategy of selecting an item from a cued category that rhymes with the irrelevant speech (e.g. what is the ANIMAL that rhymes with pog-vog-yog) would produce high levels of target recall when the speech rhymed with the target and high levels of block-1 recall when the speech rhymed with the foil. This is exactly the same pattern predicted by the feature interaction assumption.

Consequently, the following experiments are all designed to minimize the likelihood that such a strategy would be used.

It is likely that strategy use could be reduced if the irrelevant speech rhymed with a list item on only a few trials and not on the majority of trials as described above. Thus in the current experiment the auditory material only rhymed with the foil on interference trials. This represented less than a quarter of all trials in the experiment. On the other trials the irrelevant speech bore no relationship to any of the list items. If the feature interaction perspective is correct then the primary influence of the irrelevant speech should be observed in an increased number of block-1 intrusions on these trials.

Methods

Participants

Twenty students participated in this experiment.

Materials

The experiment consisted of forty-two trials of which twelve were one-block trials and thirty two-block trials. The thirty critical trials consisted of ten non-interference trials and ten interference trials where the irrelevant speech was unrelated to either the target or the foil and a set of ten interference trials where the irrelevant speech rhymed with the foil. In other words, the irrelevant speech material rhymed with a list item on only 10 of the 42 trials in the experiment. The materials used on the study lists and the procedures used to create the experimental trials were identical to those used in Experiment 2. That is, for each subject thirty of the forty categories were randomly assigned to the three conditions. In the case of the no interference trials, the target was placed among three filler items in the second block of each trial. The first block consisted of four filler items. For the 20 interference trials, the target and the foil were situated among three filler items in their respective blocks. Again unique sets of materials were created for each subject and the order of presentation of the trials was randomised for each subject as well.

The change in this experiment reflected the selection of the irrelevant speech material. Three non-words were generated for each foil, yielding a total of forty sets of irrelevant non-words. Each set of non-words rhymed with one another and with the corresponding foil. For example, the irrelevant speech for the item *shower* was *fower-dower-gower*. In the case of the rhyming interference trials, the non-words that served as the irrelevant speech rhymed with the foil. In the case of the no-interference and standard interference trials, the non-words from the remaining categories were randomly allocated to the different trials. In this way, the non-words did not rhyme with either the target or the foil.

Procedure

The procedures were similar to those used in Experiment 1 where irrelevant speech was presented during the study phase and recall was tested immediately.

Results

Correct recall The data summarised in Figure 3 were analysed by a one-way repeated measures analysis of variance. There was no significant difference in target recall between the three conditions, F (2, 19) = 1.42, MSe = 0.02.

Block-one intrusions As evidenced in Figure 3 participants made significantly more block-1 intrusion when the irrelevant speech rhymed with the foil than when the irrelevant speech had nothing in common with the list items, t (19) = 2.25.

Discussion

The interference effects that were observed in Experiment 2 appear to be replicated in the current experiment but are somewhat weaker. As was the case in Experiment 2, target recall differences between the no-interference and standard interference effects are minimal. More omission errors are made in the no-interference condition than in the standard interference condition, but there are a substantial number of block-1 intrusions. In fact the level of block-1 intrusions across the two experiments is almost equivalent.

The effects of between-stream similarity are evident in the data. Although there are minimal differences in target recall for the three conditions, there are significant differences in the pattern of errors made. Omission errors are relatively frequent in the no-interference and standard interference conditions and block-1 intrusions are less common. However, when the speech rhymes with the foil, omissions decrease and block-1 intrusions increase. The difference in the number of block-1 intrusions between the standard interference and the plus-rhyme interference condition provides initial evidence for between-stream similarity effects. That is, the presence of rhymes in the irrelevant stream appears to strengthen the foil such that it is more likely to be mistakenly recalled as the target item. We think that this pattern is not likely to reflect strategy usage but this issue will be again addressed in subsequent experiments.

Experiment 4

The use of rhymes in the irrelevant speech is one means of testing the feature interaction assumptions. A stronger test of these assumptions involves distributing features across different items. Target recall should be enhanced when the irrelevant features support the target; target recall should decrease and foil recall increase when the irrelevant features support the foil.

Distributing the relevant phonemic information across items in the irrelevant stream should make it more difficult for people to use this information in a strategic manner to predict the relevant list item. We also increased the number of items in the auditory stream and altered their nature as well. The auditory stream in this experiment consisted of 20 words which were either filler items from the current trial, fillers from previous trials in the experiment, some new words and of course some where the critical items that contained the relevant phonemes. If people did adopt a strategic approach to the task, we hoped that they would focus on an “old words”/ “new words” dimension rather than on the phonemic characteristics of the speech.

Secondly, the second block of items was increased from four words to five words. If participants are not using the irrelevant speech to predict the target then levels of recall performance should be lowered in comparison to previous experiments. If they are using the irrelevant speech, absolute levels of performance should not change appreciably from those found in the previous experiments. Thirdly, the effects of irrelevant speech were investigated under conditions where the recall task was made more difficult by inserting a one-second filled retention interval before recall. It was assumed that the distractor activity would increase the level of task difficulty and again produce a decrement if subjects were not using the irrelevant speech.

Methods

Participants

Forty students participated in this experiment. Group allocation was done according to time of arrival. For the first twenty participants recall was required immediately after presentation of the final item while the second group of twenty participants was required to recall the relevant item after a one-second filled retention interval.

Materials

The experiment consisted of forty two-block and twelve one-block trials. The materials and methods used in the construction of the trials were those used in Experiment 3. The only difference was that the number of items in the second block increased from four to five words with the addition of another filler item (the pool of filler items were increased by forty items derived from the same sources as the original pool of filler items); the target item still appeared in the second or third serial position. On ten interference trials and ten non-interference trials the irrelevant speech supported the target item. While on a further ten interference trials the irrelevant speech supported the foil. However, for ten non-interference trials the irrelevant speech supported the non-presented foil.

Significant changes were made to the irrelevant speech. For each of the two-block trials in this experiment the irrelevant speech consisted of twenty meaningful words. The final five words within the set of twenty irrelevant speech items were critical to the rhyme manipulation. Five words were generated for each foil and target item. These words did not have category membership in common with the foil and target items. Two of the five words had the same word stem as the category instance, another two words shared the same word ending as the item and the remaining word shared the same consonant but differed in the vowel used. Thus, the initial consonant, vowel and terminal consonant appeared more than once through out the presentation of the five irrelevant speech items. The five irrelevant speech items were presented in the order of stem, ending, stem, ending, and consonant, for example, the five irrelevant speech words and order of presentation of irrelevant speech for the target item DOVE were **du**mp l**ove**, **du**st, gl**ove, d**i**v**e.

The remaining fifteen items were a mixture of the filler items that appeared in the trial being viewed by the participant, items that had appeared on earlier trials and novel items. Therefore, some of the filler items viewed by participants in the first block of words were repeated as irrelevant speech during the second block of words and some of the filler items of the second block of words were articulated as irrelevant speech during the first block of items. These 15 items were always the first items to be presented in the irrelevant stream. These were followed by the five items that reinforced either the target or the foil.

Eight irrelevant speech items were generated for the one-block trials. The irrelevant speech did not support any of the four items that comprised the one-block trials.

Unique sets of trials with unique sets of irrelevant speech material were generated for each subject and the order of the trials was randomised for each subject.

Procedure

The presentation conditions were much the same as those used in the previous experiments, save for every word that the subject studied in the bottom left of the screen, two words appeared in the top right hand corner and were articulated by the experimenter. That is, the words in the study list were presented at a rate of 1 word per second and the words in the auditory stream were presented at a rate of 2 words per second. This ensured that the irrelevant items in the auditory stream were always heard after the foil and/or the target had been seen.

The only other change was that while one group was tested immediately the other group was tested after a 1-second retention interval. For this group a two-digit number appeared in the bottom left corner of the computer screen immediately after the final list item had been presented. Participants were required to indicate whether the number was larger or smaller than fifty by pointing their finger up for larger or down for smaller. The category cue appeared one-second later prompting recall of the target item.

Results

Correct recall A 2 x 2 x 2 mixed design analysis of variance was conducted on the correct recall data. The analysis confirmed what is evident in Figure 4. No significant main effect was evident for the interference manipulation, F (1, 38) = .06, MSe = 0.0, verifying that the levels of correct recall are the same for interference and non-interference conditions. The main effect for retention interval was not significant, F (1, 38) = .49, MSe = 0.12. A significant main effect for between-stream similarity emerged, F (1, 38) = 50.58, MSe = 0.02, confirming that target recall was significantly reduced when the irrelevant speech supported the foil than when it supported the target item. The analysis produced a significant similarity by retention interval interaction, F (1, 38) = 5.92, MSe = 0.02. Simple effects analyses showed that the between-stream similarity effect was significant for both immediate, F (1, 19) = 12.60, MSe = 0.02, and delayed recall, F (1, 19) = 40.25, MSe = 0.03, but the effect was stronger under delayed recall conditions. No other interactions emerged as significant.

Block-one intrusions The mean probability of block-one intrusions for each condition is also presented in Figure 4. More block-one intrusions occurred when the irrelevant speech supported the foil than when the target was supported, F (1, 38) = 36.33, MSe = 0.01; more block-one intrusions were produced under delayed recall than immediate recall, F = (1, 38) = 9.73, MSe = 0.01. Similarity and retention interval interacted significantly, F (1, 38) = 7.98, MSe = 0.01. Simple effects revealed that the similarity effect was significant for both immediate recall, F (1, 19) = 6.16, MSe = 0.01, and delayed recall, F (1, 19) = 33.55, MSe = 0.01; the effect was greater under delayed conditions.

Discussion

The results replicated the basic characteristics of Experiments 3. Although interference effects are not evident with target recall the same error patterns as Experiments 3 have emerged. Irrelevant speech that reinforces the target item produces high levels of target recall and low levels of block-one intrusions and omission errors. However when the last five items of irrelevant speech shared features with the foil, target recall decreased in frequency and production of the foil and omission errors increased. Immediate and delayed recall produced the same patterns of performance but the effects of irrelevant speech were greater after a filled retention interval of one-second. In short, the data suggest that between-stream similarity is having an impact upon cued recall performance.

With respect to the strategy issues, the changes in methodology have reduced the levels of performance quite substantially compared to Experiment 3, without affecting the pattern of performance in any way. There is no reason to expect this change if subjects were using the irrelevant speech to identify the target item. Therefore, strategy effects do not appear to be the prime determinant of recall performance.

However, there are still some potentially problematic aspects of our experimental procedure. Firstly the experimenter, who is familiar with the experimental manipulations and hypotheses, articulated the irrelevant speech thus introducing a possible bias. In the next experiment, the irrelevant speech is produced by a person who is blind to the intent of the experiment. Secondly, the phonemic features of either the target or foil were distributed across the final five irrelevant speech items. These final five items were the last words the participants heard before recall. Some might argue that the effects are due to the effects of echoic memory. If so, placing the five critical items earlier in the list should attenuate the similarity effect. Thus, in the next experiment, the five words are presented together as the fourteenth to eighteen words in the stream, with the final two words being unrelated to the foil or the target, but acting as a stimulus suffix.

Experiment 5

Method

Participants

Forty participants took part in this experiment and they participated in pairs. One member of each pair served as experimental subject and the other read the irrelevant speech items as they appeared on the screen.

Materials

The materials were identical to Experiment 5.

Procedure

The procedure was identical to Experiment 5 except for three features. Firstly, only the delayed recall version was used; secondly, the irrelevant speech was read by one of the experimental subjects rather than the experimenter, and thirdly, the critical irrelevant speech items appeared not at the end of the list, but were followed by two items.

Results and Discussion

Correct recall The target recall data, displayed in Figure 5 was submitted to a 2 x 2 repeated measures analysis of variance. There was no significant difference in target recall between interference and non-interference trials, F (1, 19) = .07, MSe = 0.04. A main effect for between-stream similarity emerged, F (1, 19) = 4.36, MSe = 0.02, confirming that target recall is significantly reduced when the irrelevant speech contained the phonemic features of the foil. The interference by irrelevant speech interaction did not reach significance.

Block-one intrusions Planned comparisons confirmed the pattern in Figure 5 in that significantly more block-one intrusions were produced when the irrelevant speech supported the foil than when the irrelevant speech reinforced the phonemic characteristics of the target item, t (19) = 3.68.

Between-stream similarity effects are again present in the current data. Both target recall and block-1 intrusions are influenced by the items that are being presented in the irrelevant stream. Consequently, whether or not the person producing the auditory material is the experimenter or not seems to have little bearing on the outcomes of the experiment. Moreover, whether or not the critical words appear as the last items in the list also seems to be relatively trivial. Echoic memory does not play any significant role in the outcome.

Experiment 6

In most previous studies of irrelevant speech the speech has been presented via an audio tape or via digitised speech on computer. Our final experiment uses digitised speech to present the irrelevant material. Furthermore, in this experiment all subjects see the same set of trials. Each trial is associated with two versions of the words in the irrelevant stream. Each of these streams contains 18 words, and across versions each stream is identical save for five words that differ. In one version the five words support the target and in the other stream the five words support the foil. To the extent that having a live body produce the irrelevant speech is not a contributing factor to our results, and that the results are due to the phonological overlap between visual and auditory stream, we expect to see the same pattern that has been found in the previous experiments. If we find such an effect we can say with some certainty that the effect is due only to the five words in the irrelevant stream because all other factors are identical.

Method

Subjects

Twenty students studying psychology at the Australian Catholic University participated for course credit.

Materials

All subjects saw the same set of trials. One of the sets that had been used in Experiment 5 was randomly selected. The trials consisted of 12 one-block trials, 20 no-interference and 20 interference trials all in a random order. The irrelevant speech that had been associated with these trials in Experiment 5 was also used. In Experiment 5, the auditory stream for each trial consisted of 20 words, of which 15 were filler items and five items shared phonological features with either the target or the foil. These five words appeared in the 14^{th} to 18^{th} position in each stream. In creating the irrelevant speech for this experiment, the auditory stream on each trial was digitised and stored on computer. This served as version one of the irrelevant speech. On half the trials the five critical items shared features with the target and half shared features with the foil. The second version was created by replacing the five words in positions 14 to 18 with the alternative words. Thus if the irrelevant items supported the target in version 1, in version 2 they supported the foil. Likewise if the irrelevant items supported the foil in version 1, they supported the target in version 2. Thus, all subjects saw the same trials and on each trial heard the same irrelevant speech, save for five words in the irrelevant stream that differed across the two versions. Half the subjects were given version 1 of the irrelevant speech and half heard version 2.

After the lists of digitised speech had been created, we discovered that each sound file ran for 12 seconds, which meant that the speech was finishing as the category cue appeared on the screen. To ensure that the irrelevant speech ended at the same time that the final list word disappeared from the screen, the speech files were edited to 10 seconds worth of speech. This was done by removing the first two or three filler items from each speech file. Importantly, the five critical items were not paired with the presentation of either the target or the foil.

Procedure.

The procedure was identical to that used in Experiment 5, save that the irrelevant speech was presented over the computer speakers.

Results and Discussion

Correct recall The target recall data, displayed in Figure 6 was submitted to a 2 x 2 repeated measures analysis of variance. There was a significant difference in target recall between interference and non-interference trials, F (1, 19) = 22.63, MSe = 0.01. A main effect for between-stream similarity emerged, F (1, 19) = 33.09, MSe = 0.04, confirming that target recall is significantly reduced when the irrelevant speech contained the phonemic features of the foil. The interference by irrelevant speech interaction did not reach significance.

Block-one intrusions Planned comparisons confirmed that significantly more block-one intrusions were produced when the irrelevant speech supported the foil, t (19) = 6.45.

Between-stream similarity effects are again present in the current data. The items that are being presented in the irrelevant stream influence both target recall and block-1 intrusions. The effects are localised to five words that are appearing in the irrelevant stream. In all other aspects of the experiment all subjects are experiencing the same material on the study lists and in the irrelevant stream. Whether the speech is digitised or spoken by a second person in the room also seems to be irrelevant to producing the effect.

General Discussion

There are two sources of evidence indicating the presence of between-stream similarity effects in the current data. The first is that target recall differs depending upon the items that are in the auditory stream. When the phonological characteristics of the speech support the target item recall of the target is protected on both interference and no-interference trials compared to when the speech supports the block-1 foil. The second source is more direct. When the phonological features of the irrelevant speech are common with the phonological features of the foil then increased recall of the foil results. In short, we can influence the probability that the target or the foil will be recalled by varying the phonological characteristics of the items in the irrelevant stream. If between-stream similarity were irrelevant this pattern would not emerge.

The effects of irrelevant speech on target and foil recall are observed across a variety of experimental conditions. The effects are present on an immediate test and a delayed test. They are present on short (4 word) and longer (5 word) lists. They are present when the irrelevant speech consists of steady-state speech (Experiment 3) and changing-state speech (Experiments 4-6). The effects are replicable and do not seem to be attributable to students adopting a strategy of using the irrelevant speech to predict the list item.

While between-stream effects are observed they are potentially problematic. The results indicate that shared features strengthen the representations of an item, be it the target or the foil, instead of producing interference. This seems to be at odds with the immediate serial recall data in which irrelevant speech always has a detrimental effect, and between-stream similarity has no effect. It might be the case that irrelevant speech effects might reflect different processes when it comes to the two tasks. That is, the cued recall results may have no relevance to the immediate serial recall results or to current theories of the irrelevant speech effect. In the next section we look at the results in the light of what is known about serial recall in an attempt to address this possibility.

Cued Recall Versus Serial Recall

The results are consistent with some aspects of what is know of irrelevant speech effects in immediate serial recall and inconsistent with others. Experiments 1 and 2 indicated that the cued recall task was hurt by irrelevant speech, as is performance in immediate serial recall. This difference was reflected primarily in increased numbers of omission errors.

As is the case with serial recall, there were indications that changing-state irrelevant speech produced more disruption than steady-state speech, although the effects were not as robust as those observed in the Jones and Macken (1995) experiments. This may reflect a cued recall versus serial recall difference, or it may be the case that the Jones and Macken (1995) data were more pronounced that is normally the case. Thus, while strong steady-state changing-state differences were found in the Jones and Macken (1995) data, much weaker effects were observed in the Larsen, Baddeley and Andrade (2000) research, which also involved immediate serial recall. In fact, the current cued recall results look a lot like the Larsen et al. results in that in both cases the means are in the expected direction, but differences are not statistically reliable.

There are other possible explanations for the weak differences between changing and steady-state speech in the cued recall data. While we designed the two types of irrelevant speech to be instances of changing-state and steady-state, the two types may have ended up being functionally equivalent, particularly in the case of where three non-words were presented during study. Effectively students may have perceived the repetition of a single, long nonsense word rather than the repetition of three distinct non-words for both steady-state and changing-state materials. Such an account might also explain why irrelevant speech effects were stronger when irrelevant speech was presented during a brief rehearsal period than when it was presented during study, many repetitions are likely to produce perception of a single long non-word that three one syllable non-words.

The one obvious inconsistency between cued recall and serial recall is that between-stream similarity effects were present in the cued recall task but not readily observable in immediate serial recall (Bridges & Jones, 1996; Jones & Macken, 1995; Larsen et al., 2000; LeCompte & Shaibe, 1997). One reason for the discrepancy might be that we have relied upon a PI task to show the effects. In the standard immediate serial recall tasks PI is normally not manipulated and consequently there is little opportunity for intrusions to take place. Our data however, do suggest that we have a task that is sensitive to between-stream similarity effects and there is no reason why the cued recall task could not be adapted for serial recall. If cued recall and immediate serial recall both rely upon phonological codes then between-stream effects might be observed if a PI version of immediate serial recall were employed.

The use of a PI task also provides another possible explanation for the absence of effects in immediate serial recall. When target recall in the cued recall task is considered there does not seem to be a lot of evidence for any effects of PI. The effects we are observing tend to emerge in differential omissions and block-1 intrusions. This finding has potential implications for between-stream similarity effects in serial recall, particularly when it comes to item and order errors. Beaman and Jones (1998) have shown that irrelevant speech increases the likelihood of both item and order errors in immediate serial recall. However, in the studies that have explored between-stream similarity effects in immediate serial recall (Bridges & Jones, 1996; Jones & Macken, 1995; Larsen et al., 2000; LeCompte & Shaibe, 1997) no analyses of the errors have been made. It is possible that between-stream similarity might have opposite effects upon item and order information such that they cancel each other out. That is, it is possible that similar irrelevant speech might produce more order errors in immediate serial recall, but could produced enhanced item information along the lines observed in the current experiments. The facilitative and interfering effects could well cancel each other out to give the impression that between-stream similarity had no effect. Fallon, Groves & Tehan (1999) report such an effect with phonological similarity effects. On a delayed test target recall for phonologically similar and dissimilar lists were equivalent suggesting that phonological codes had dissipated. However, an error analysis showed that item information was better for the phonologically similar lists but order errors were more pronounced. Contrary to the conclusion suggested by target recall, the error analyses indicated that phonological codes were still playing a role in the task.

Although general opinion assumes that between-stream effects are not present in immediate serial recall, the studied involved only a limited set of conditions; conditions that are probably less than optimal for showing between-stream effects. The memory materials used in the serial recall studies have always been digits (Bridges & Jones, 1996; LeCompte & Shaibe, 1997) or letters (Jones & Macken, 1995; Larsen et al., 2000) and relatively long lists have been employed. It is known that serial recall performance is sensitive to whether a closed pool items is used rather than an open pool (Coltheart 1993; LaPointe & Engle, 1990; Nairne & Kelley, 1999), and there are clear differences in aspects of serial recall when words are used as the memory stimuli as opposed to digits or letters. Furthermore, with long lists subjects tend to abandon phonological coding (Salamé & Baddeley, 1986). Under such circumstances between-stream similarity effects are less likely to be expected. In short, there is still much that is not known about between-stream similarity effects in serial recall. The current results suggest that between-stream similarity effects might be possible in serial recall if the study lists consisted of short lists of words that came from a large word pool and if alternative scoring procedures were adopted.

Another reason for why between stream effects have not been observed in serial recall is that similar irrelevant speech may actually support recall of the list items rather than produce a source of interference as is commonly assumed. The conditions used in the serial recall research are the same as those used in the current experiments where the irrelevant speech supported the target item. Consequently, similar irrelevant speech may tend to enhance recall of the list items. However, any enhancement effects would critically depend upon the level of degradation of the phonological features of the list items. If there was very little degradation, then feature adoption from the irrelevant stream would have minimal effects, in that intact features of the item would be replaced by the same feature from the irrelevant stream, producing a zero net effect. A zero net effect might also result if facilitative effects offset interfering effects of feature adoption. Again these ideas are testable.

Of course there are other possible reasons for a difference between cued recall and serial recall results. Clearly, there are a number of differences between the experimental procedures. For a start the current research used cued recall where a single item response is made under conditions where order memory is not required. The serial order research involves multiple item responses in a situation where order information is critical. Either of these differences could be crucial. In fact, Jones (1993) has always argued that the changing-state advantage should be restricted to tasks that involve retention of order. The fact that differences between changing-state and steady-state speech are much weaker in the cued recall task is consistent with Jones' view. Consequently, one might argue that because the tasks are to some extent different, it is possible that between-stream effects might be found in one task but not the other. One way that this might happen is if the effects of between-stream similarity in immediate serial recall are masked by other sources of forgetting or interference that are occurring during the recall process itself. It is now clear that serial recall is more complicated than simply dumping the contents of a short-term buffer. Instead it seems to involve repeated retrievals, possible multiple searches and redintegration of degraded traces (Brown & Hulme, 1995; Cowan et al, 1992). Other sources of interference may be more robust and mask the effects of between-stream similarity in the same way that articulatory suppression masks the effects of irrelevant speech (Neath, 1999).

One final consideration concerns the temporal distribution of the list items and the irrelevant stream. There was clear evidence for enhanced recall of an item early in the list (the foil) when the similar material occurred late in the irrelevant stream. In the absence of appropriate controls in the current experiments we cannot be certain whether the between-stream similarity effects on target recall were facilitative or had little impact at all. It is quite plausible that phonological representations of the latter list items are relatively intact and gain little additional support from items in the irrelevant stream. Furthermore, we have not explored what happens when the similar items are presented early in the irrelevant stream, or are evenly distributed throughout the stream. Clearly the temporal characteristics of the irrelevant item could be crucial to finding between-stream effects. The current research, however, does suggest that if cued recall and serial recall are supported by the same processes and codes, then between-stream similarity effects could be found in serial recall, at least for early serial positions, if the similar items were presented late in the auditory stream.

Implications for Models of Irrelevant Speech

The current research was motivated by the need to test feature interaction explanations of irrelevant speech effects. The results of Experiments 4 to 6 complement the Tehan and Humphreys (1998) research. That is, in both cases features of irrelevant items influence recall in the same way by enhancing the likelihood that a block-1 foil will be recalled. In the case of Tehan and Humphreys (1998) the features came from filler items in the list. In the current experiment, the features came from items in the auditory stream. However, in both instances the effects were the same; irrelevant items influenced recall. These results provide strong evidence for feature interaction.

The current results provide support for an essential assumption of Neath's (1999, 2000) adaptation of the feature model. The model assumes that features of the list item and features of the irrelevant items interact with each other. Differences in target recall and in block-1 intrusions are most readily explained in term of a memory trace that consists of the combined features of list and auditory items.

The current results, and those of Tehan and Humphreys (1998), also speak to the way in which the features interact or are absorbed. The criticisms that Baddeley (2000) and Jones and Tremblay (2000) level at Neath's (2000) model appear to assume that feature adoption should be position specific. That is, the first phoneme from an irrelevant word should be absorbed into the first phoneme position of a list word and so on. This is precisely the way we have manipulated feature interaction by ensuring that the features from one source are in the same within-word position as the item in the second source. We have shown that this does influence recall in the way that Baddeley and Jones and Tremblay assume. In short, the data support position-specific feature interaction assumptions.

While feature adoption is an essential component of the feature model, the formal implementation of the adoption process does not assume position specific interactions. In the feature model one of the items in the irrelevant stream is randomly paired with the list word and then a random half the modality independent features of the list item are replaced with a random half of the features of the auditory item. Importantly, the features of the auditory item that are adopted do not necessarily have to migrate to the same feature position in the list item (Neath, personal communication). That is, a feature that occupies position *n* in the vector representing the word in the irrelevant stream can be adopted into position *n+1, n+2, n-1, n-2,* etc. of the list item. In terms of the example we used earlier, where cat is a list item and cot is an item in the auditory stream, implementing feature adoption in a random order produces six possible short-term traces; either *cat, tac*, *oac*, *oat, tao,* or *cao*. Obviously, the degree of match between the short-term trace and the long-term trace critically depends upon the specific form of the short-term trace that emerges from the feature adoption process. Given that in the feature model, the number of modality independent features involved is 20 instead of the three that we have used in the example, the chances of obtaining a short-term trace that facilitates recall are relatively small. Randomised-position feature adoption is probably responsible for the fact that the feature model does not produce the robust between-stream similarity effects that Baddeley and Jones and Tremblay expect.

The way in which between-stream similarity is operationalised in Experiments 4 to 6 is also problematic for the way in which the feature model works. In these experiments features from three different words appear to be adopted into the short-term trace. In the feature model a single word in the irrelevant stream is paired with a list item and features from that word are adopted into the short-term trace. It is not usually the case that multiple words are paired with each list word. Feature adoption from three words could happen if the target item was rehearsed at least three times and the relevant features were adopted in each case. However, it is hard to see how the features from three different words could be adopted into the trace of the foil when it is unlikely that the foil is being rehearsed. The current results suggest that the features of the foil, the target and the features of the irrelevant words are all simultaneously active in memory at the point of recall. This is not an assumption of the feature model but it is an emergent property of some distributed memory models that employ sparse distributed representations (Chappell & Humphreys, 1994; Tehan & Fallon; 1999; Tehan & Humphreys, 1998).

Do the current findings supporting position specific interactions invalidate the feature model? The answer is clearly no. We have not tested the interfering and facilitative effects of irrelevant speech where the relevant features in the irrelevant stream appear out of position. Until this is done there is no firm evidence for how problematic the current results are for the specific implementation of the feature model. It should be said however, that Li, Schweickert and Gandour (2000) have looked at phoneme order with respect to the phonological similarity effect. They demonstrated that words like disk and skid produced a phonological similarity effect in the same way as the standard manipulation involving words like disk and desk. There is no reason to expect that results would be different if desk or skid appeared in the irrelevant stream in the current task, but it is certainly testable.

The results also have implications for the other models of irrelevant speech. The O-OER model (Jones, 1993) argues that between stream similarity effects should not affect performance because interference is determined solely by variability within-stream. The fact that we have observed between-stream effects may have implications for one particular aspect of that model, assuming that the cued recall task is one for which the O-OER model applies. We would suggest that the task in one in which the model does apply. Although the task does not require order memory, subjects are encouraged to rehearse items and the most likely form of rehearsal is rote rehearsal. The fact that irrelevant speech affects performance at all, and the fact that changing-state speech produces added disruption certainly suggests that the model is relevant to the task.

Do the results invalidate the O-OER model? Again the answer is no. There is much in the data that is totally consistent with the model. Furthermore, we have been intentionally specific with our use of the term irrelevant speech rather than irrelevant sound as a way of indicating the limitations of the current research. The interaction of speech with verbal items in memory is readily handled from a feature adoption or any other item interaction perspective. The interaction of non-speech material with verbal material in memory is less readily explained. So while between-stream similarity effects might be a potential Achilles heel for the O-OER model, irrelevant sound effects remain a potential Achilles heel for feature adoption and item interaction effects models of the type that Neath (2000) and Tehan and Humphreys (1998) describe.

The presence of between-stream similarity effects is a prediction of Baddeley's (1986) working memory model. Thus, the current results are consistent with this assumption. However, because the effects of irrelevant speech within the phonological store remain unspecified, it is impossible to say whether the results are consistent with the model. If the current results are to be explained within the phonological store framework, then the results clearly indicate that the features of the items in the store interact. That is, the resultant traces would be a combination of features from different items. It is easy to see how this would produce a deficit in recall, but it is not easy to see how a degraded trace might lead to enhanced recall of a block-1 foil in a cued recall task.

In conclusion, the first two experiments show that irrelevant speech effects can be observed in a cued recall task for which there is no obvious order requirement although subjects in some conditions are instructed to rehearse. The results are consistent with the view that irrelevant speech has an impact because it produces interference with phonological representations. Experiments 3 to 6 tested the idea that features of the irrelevant items could interact with features of the list items, a prediction that is at the centre of Neath's (1999, 2000) model of irrelevant speech effects. We produce relatively direct support for the feature adoption assumption. Feature interaction considerations lead to a novel prediction that irrelevant speech can actually facilitate recall, a prediction that is not made by any of the current models. In the cued recall task we are able to show that between-stream similarity effects are observed in the task and that we can predict when the irrelevant speech will produce interference and when it will support recall. All current models predict that irrelevant speech will only have an interfering effect upon performance. The results thus produce a challenge to all current models of irrelevant speech and suggest means by which more sensitive tests of between-stream similarity effects in serial recall might be explored.

References

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

Baddeley, A. D., & Hitch G. J. (1994). Developments in the concept of working memory. Neuropsychology, 8, 485-493.

Baddeley, A. D., & Salamé, P. (1986). The unattended speech effect: Perception or memory? Journal of Experimental Psychology: Learning, Memory and Cognition, 12, 525 – 529.

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

Beaman, C. P., & Jones, D. M. (1998). Irrelevant sound disrupts order information in free recall and serial recall. Quarterly Journal of Experimental Psychology, 51A, 615-636.

Bridges, A. M. & Jones, D. M. (1996). Word-dose in the disruption of serial recall by irrelevant speech: Phonological confusions or changing-state? Quarterly Journal of Experimental Psychology, 49A, 919-939.

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

Chappell, M. & Humphreys, M. S. (1994). An auto-associative neural network for sparse representations: Analysis and application to models of recognition and cued recall. Psychological Review, 101, 103-128.

Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. Journal of Verbal Learning and Verbal Behavior, 19, 722 –735.

Colle, H. A. & Welsh, A. (1976). Acoustic masking in primary memory. Journal of Verbal Learning and Verbal Behaviour, 15, 17-31.

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

Cowan, N., Day, L., Saults, J.S., Keller, T.A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. Journal of Memory and Language, 31, 1-17.

Fallon, A. B., Groves, K & Tehan, G. (1999). Phonological similarity and Trace Degradation in the Serial Recall Task: When CAT helps RAT, but not MAN. International Journal of Psychology, 34, 301-307.

Jones, D. M. (1993). Objects, streams and threads of auditory attention. In D. D. Baddeley & L Weiskrantz (Eds.), Attention: Selection, Awareness and Control. Oxford: Clarendon press.

Jones, D. M. & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. Journal of Experimental Psychology: Learning, Memory and Cognition, 19, 369-381.

Jones, D. M. & Macken, W. J. (1995). Phonological similarity in the irrelevant speech effect: Within- or between-stream similarity? Journal of Experimental Psychology: Learning, Memory and Cognition, 21, 103-115.

Jones, D. M. & Tremblay, S (2000). Interference in memory by process or content? A reply to Neath (2000). Pscyhonomic Bulletin and Review, 7, 550-558.

Larsen, J. D., Baddeley, A. D., & Andrade J. (2000). Phonological similarity and the irrelevant speech effect: Implications for models of short-term verbal memory. Memory, 8, 145-157.

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

LeCompte, D. C. & Shaibe, D.M. (1997). On the irrelevance of phonological similarity to the irrelevant speech effect. Quarterly Journal of Experimental Psychology, 50A, 100-118.

Li, X., Schweickert, R., & Gandour, J. (2000). The Phonological Similarity Effect in Immediate Recall: Positions of Shared Phonemes. Memory and Cognition, 28, .

McEvoy, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. American Journal of Psychology, 95, 581 - 635.

Miles, C., Jones, D. M. & Madden, C. A. (1991). Locus of the irrelevant speech effect in short-term memory. Journal of Experimental Psychology: Learning, Memory and Cognition, 17, 578-584.

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

Neath, I. (1999). Modelling the disruptive effects of irrelevant speech on order information. International Journal of Psychology, 34, 410-418.

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

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

Salamé, P. & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. Bulletin of the Psychonomic Society, 24, 263-265.

Shapiro, S, I., & Palermo, D. S. (1970). Conceptual organization and class membership: Normative data for representations of 100 categories. Psychonomic Monograph Supplements, 3, 107 - 127.

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

Tehan, G., & Humphreys, M. S. (1995). Transient phonemic codes and immunity to proactive interference. Memory and Cognition, 23, 181-191.

Tehan, G., & Humphreys, M. S. (1996). Cuing effects in short-term recall. Memory and Cognition, 24, 719-732.

Tehan, G., & Humphreys, M. S. (1998). Creating proactive interference in immediate recall: Building a DOG from a DART, a MOP, and a FIG. Memory and Cognition, 26, 477-489.

Tolan, G. A. & Tehan, G. (1999). Determinants of Short-term Forgetting: Decay, Retroactive Interference or Proactive Interference? International Journal of Psychology, 34, 285-292.

Walling, J. R., McEvoy, C. L., Oth, J. E. & Nelson, D. L. Rhyme category norms. Unpublished manuscript. The University of South Florida.

Table 1.

The effect sizes between the quiet and steady-state conditions and between quiet and changing-state conditions for Experiment 1 and 2

	Quiet vs Steady-state	Quiet vs Changing-state
	Correct recall	
Experiment 1	η^2	η^2
Input Phase	.22	.40
Output Phase	.57	.72
Experiment 2		
Interference	.16	.30
Non-interference	.06	.11
	Omission errors	
Experiment 1	η^2	η^2
Input Phase	.18	.30
Output Phase	.58	.70
Experiment 2		
Interference	.05	.21
Non-interference	.02	.17
	Block-one intrusions	
Experiment 2	η^2	η^2
	.11	.14

Author Information

Correspondence regarding this article can be addressed to either author by e-mail (a.tolan@mcauley.acu.edu.au or tehan@usq.edu.au). The second author would like to thank Jim Nairne and the staff of Purdue University for their support during his sabbatical when this manuscript was developed. We would also like to thank James Nairne, Doug Nelson, Philip Beaman and an anonymous reviewer for helpful comments upon an earlier draft.

Figure Captions

Figure 1. The mean proportion of items correctly recalled and omissions made as a function of irrelevant speech and the point of presentation of the irrelevant speech. Error bars depict standard error.

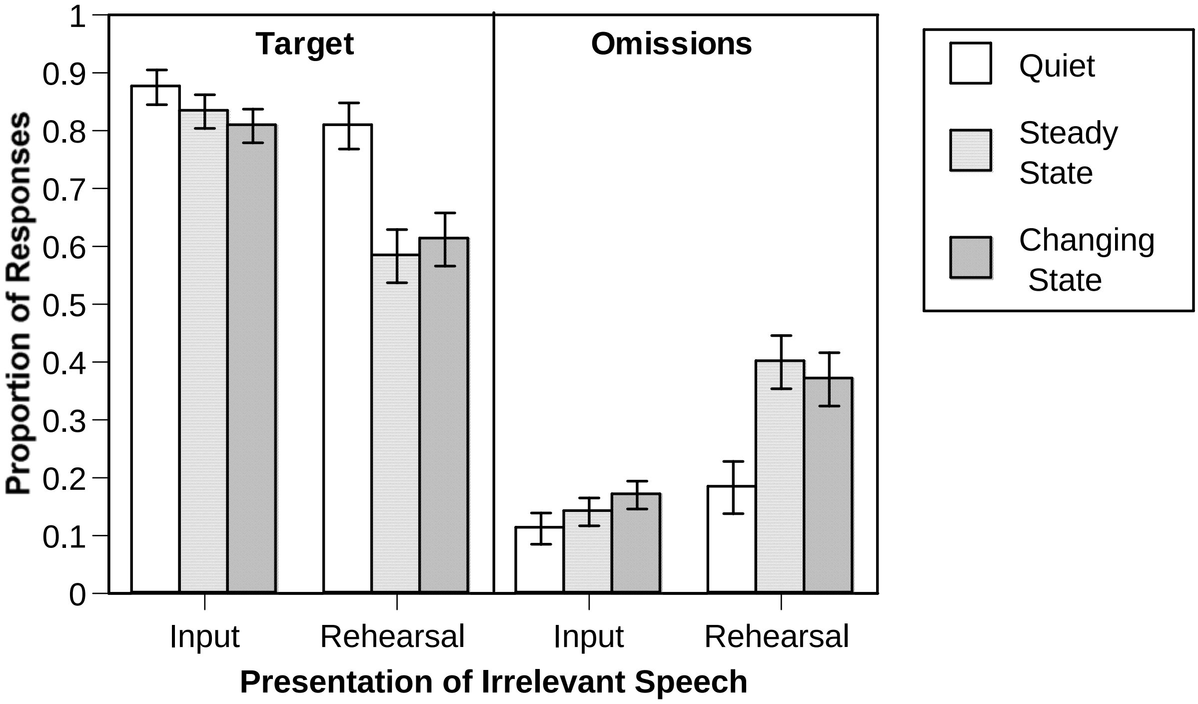
Figure 2. The mean proportion of correct recall, omission errors and block-1 intrusions as a function of interference and irrelevant speech. Error bars depict standard error.

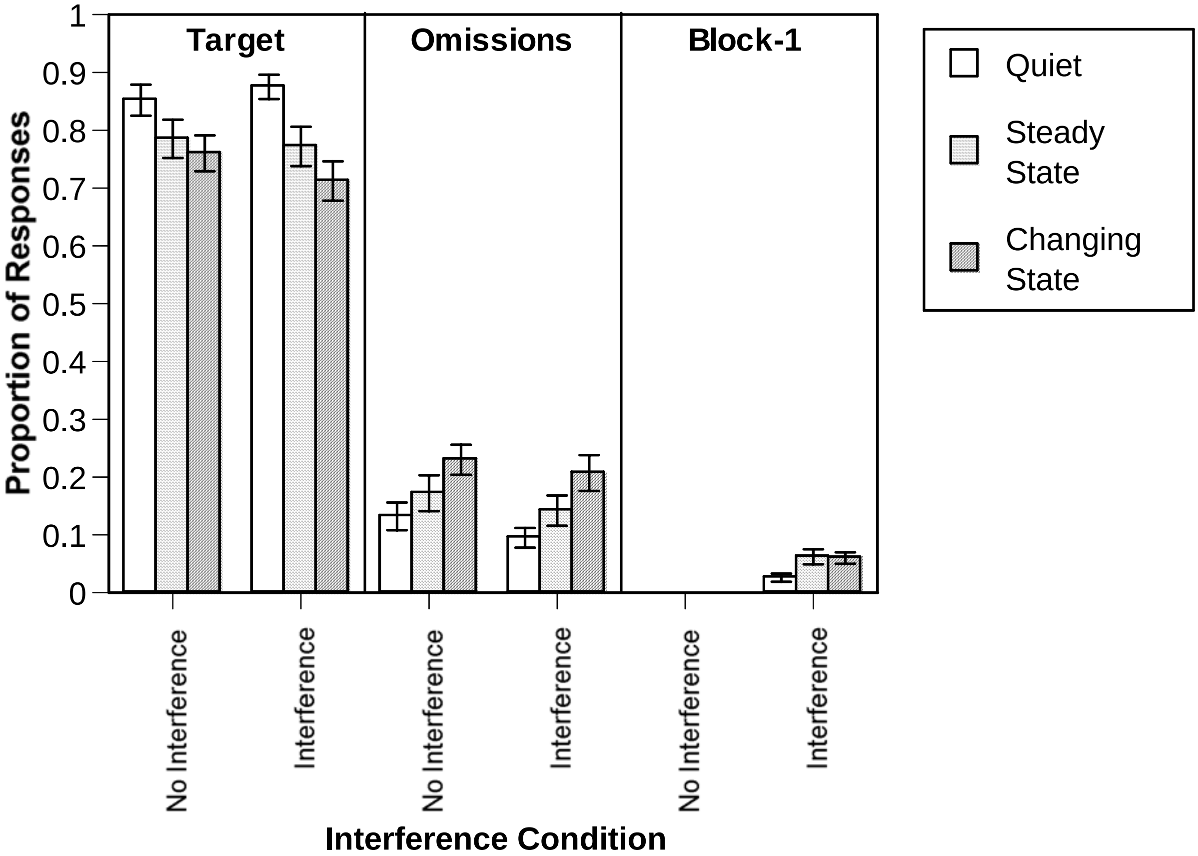
Figure 3. The mean proportion of correct recall, omission errors and block-1 intrusions as a function of interference and irrelevant speech. Error bars depict standard error.

Figure 4. The mean proportion of correct recall, omission errors and block-1 intrusions as a function of interference, irrelevant speech and retention interval. Error bars depict standard error.

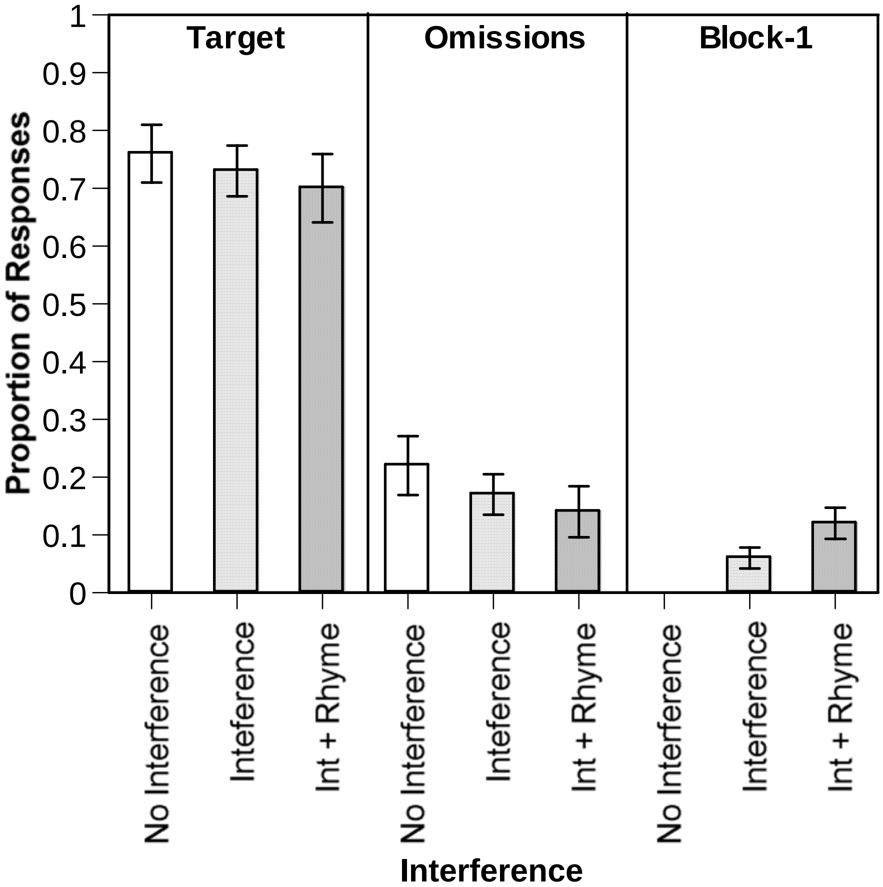
Figure 5. The mean proportion of correct recall, omission errors and block-1 intrusions as a function of interference and irrelevant speech. Error bars depict standard error.

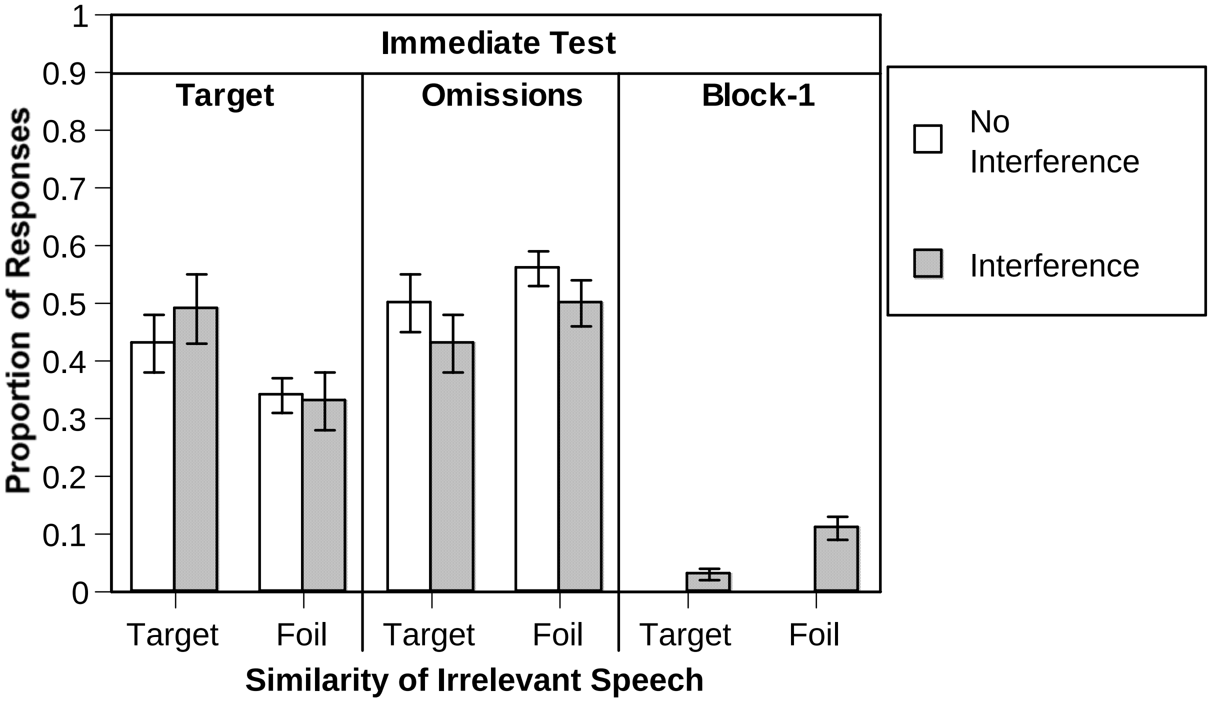
Figure 6. The mean proportion of correct recall, omission errors and block-1 intrusions as a function of interference and irrelevant speech. Error bars depict standard error.

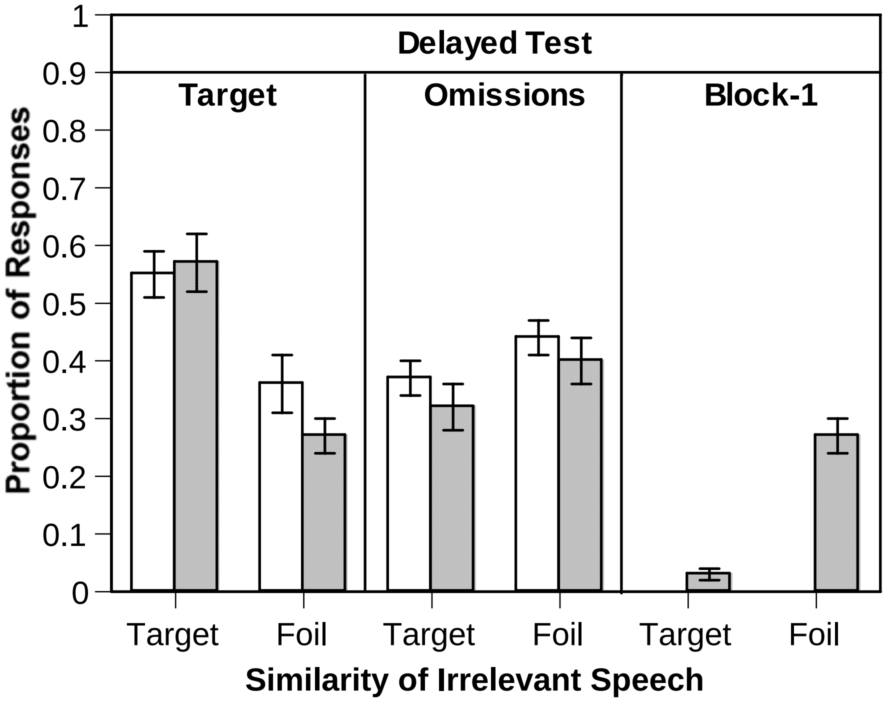

Target
Omissions
Proportion of Responses
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Input
Rehearsal
Input
Rehearsal
Presentation of Irrelevant Speech
Quiet
Steady State
Changing State

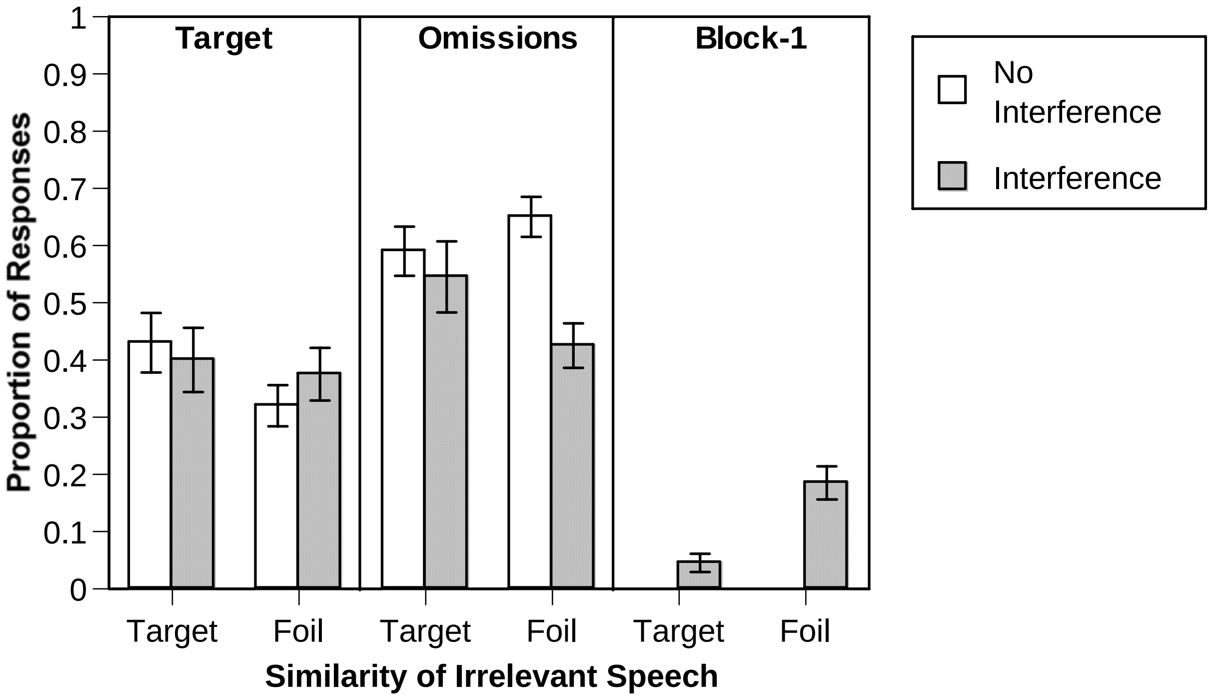


Proportion of Responses
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Target
Omissions
Block-1
No Interference
Interference
No Interference
Interference
No Interference
Interference
Interference Condition
Quiet
Steady State
Changing State

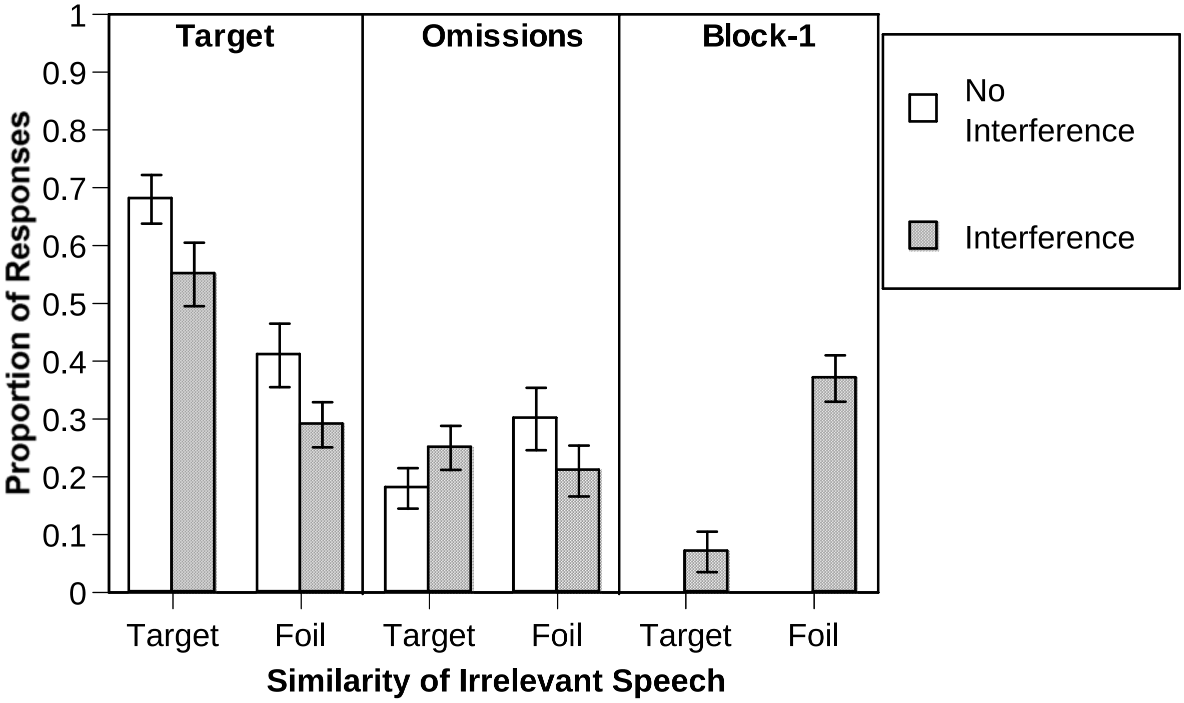

Target
Omissions
Block-1
Proportion of Responses
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
No Interference
Inteference
Int + Rhyme
No Interference
Inteference
Int + Rhyme
No Interference
Interference
Int + Rhyme
Interference


Immediate Test
Target
Omissions
Block-1
No Interference
Interference
Proportion of Responses
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Target
Foil
Target
Foil
Target
Foil
Similarity of Irrelevant Speech


Delayed Test
Target
Omissions
Block-1
Proportion of Responses
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Target
Foil
Target
Foil
Target
Foil
Similarity of Irrelevant Speech



Proportion of Responses
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Target
Omissions
Block-1
Target
Foil
Target
Foil
Target
Foil
Similarity of Irrelevant Speech
No Interference
Interference



Target
Omissions
Block-1
No Interference
Interference
Proportion of Responses
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Target
Foil
Target
Foil
Target
Foil
Similarity of Irrelevant Speech