Central Lancashire Online Knowledge (CLoK)

Title	Boundaries of Semantic Distraction: Dominance and Lexicality Act at Retrieval
Type	Article
URL	https://clok.uclan.ac.uk/13935/
DOI	https://doi.org/10.3758/s13421-014-0438-6
Date	2014
Citation	Marsh, John Everett, Perham, Nicholas R., Sorqvist, Patrik and Jones, Dylan M. (2014) Boundaries of Semantic Distraction: Dominance and Lexicality Act at Retrieval. Memory & Cognition, 42 (8). pp. 1285-1301. ISSN 0090-502X
Creators	Marsh, John Everett, Perham, Nicholas R., Sorqvist, Patrik and Jones, Dylan M.

It is advisable to refer to the publisher's version if you intend to cite from the work.
https://doi.org/10.3758/s13421-014-0438-6

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

Boundaries of Semantic Distraction: Dominance and Lexicality Act at Retrieval

John E. Marsh[1,2], Nick Perham[3], Patrik Sörqvist[2,4], and Dylan M. Jones[5]

[1]School of Psychology, University of Central Lancashire, Preston, UK.

[2]Department of Building, Energy and Environmental Engineering, University of Gävle, Gävle, Sweden.

[3]Department of Applied Psychology, Cardiff Metropolitan University, Cardiff, UK.

[4]Linneaus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden.

[5]School of Psychology, Cardiff University, Cardiff, UK.

RUNNING HEAD: Boundaries of semantic distraction

Correspondence: John Marsh, School of Psychology,

University of Central Lancashire, Darwin Building, Preston, Lancashire,

United Kingdom, PR1 2HE.

Phone: (+44) 17 72 893754

email: jemarsh@uclan.ac.uk

Boundaries of Semantic Distraction: Dominance and Lexicality Act at Retrieval

Abstract

Three experiments investigated memory for semantic information with the goal of determining boundary conditions for the manifestation of semantic auditory distraction. Irrelevant speech disrupted the free recall of semantic category-exemplars to an equal degree regardless of whether the speech coincided with presentation or test phases of the task (Experiment 1) and occurred regardless of whether it comprised random words or coherent sentences (Experiment 2). The effects of background speech were greater when the irrelevant speech was semantically related to the to-be-remembered material, but only when the irrelevant words were high in output dominance (Experiment 3). The implications of these findings in relation to the processing of task material and the processing of background speech is discussed.

Keywords: Semantic Auditory Distraction, Selective Attention, Interference-by-process, Semantic-Category Clustering

There is something special about to-be-ignored background speech. For instance, natural speech is more detrimental to reading comprehension than reversed speech (Oswald, Tremblay, & Jones, 2000) and natural speech is more detrimental to memory of prose than aircraft noise (Sörqvist, 2010). Moreover, background conversation is often a source of complaint, annoyance and self-reported distraction (Evans & Johnson, 2000; Sundstrom, Town, Rice, Osborn, & Brill, 1994; Witterseh, Wyon, & Clausen, 2004). With the exception of its effect on serial short-term memory (Buchner, Irmen, & Erdfelder, 1996; Jones & Macken, 1993; Jones, Madden, & Miles, 1992; Jones, Miles, & Page, 1990; Marsh, Hughes, & Jones, 2008, 2009), speech is invariably the most distracting type of background sound (Marsh & Jones, 2010; Sörqvist, 2010), especially when the task entails processing of meaning (Beaman, 2004; Bell, Buchner, & Mund, 2008; Jones et al., 1990; Jones, Marsh, & Hughes, 2012; Marsh, Hughes, & Jones, 2008, 2009; Marsh, Beaman, Hughes, & Jones, 2012; Martin et al., 1988; Neely & LeCompte, 1999; Oswald, Tremblay, & Jones, 2000; Sörqvist, Marsh, & Jahncke, 2010; Sörqvist, Nöstl, & Halin, 2012). For instance, recall of a visually-presented list of items taken from the same semantic category (e.g., Fruit) is more impaired when categorically related items (e.g., other Fruit) are spoken in the background in comparison with categorically unrelated items (e.g., Tools). This between-sequence semantic similarity effect is easily replicable (e.g., Marsh et al., 2008; Sörqvist et al., 2010). However, it appears to necessitate a task instruction that permits and emphasises processing of meaning. The effect is readily observed when the requirement is to free recall to-be-remembered items but fails to materialise when the task instruction emphasises recall in serial order whereby, in contrast to free recall, the processing of meaning plays a much more

subservient role (Marsh et al., 2008; see also Marsh, Vachon, & Jones, 2008). Although

effects of the semantic similarity between the to-be-ignored background speech and the

to-be-remembered material have been found in the context of a serial recall instruction

(Bell, Mund, & Buchner, 2011; Neely & LeCompte, 1999), such as a greater number of

intrusions from related than from unrelated speech into the recall protocol (Marsh et al.,

2008), a serial recall instruction typically removes the difference between related and

unrelated background speech (e.g., Marsh et al., 2008, 2009; Sörqvist et al., 2010). Thus,

one mechanism that appears to determine the degree of disruption attributable to the

semantic meaning of the background speech is the nature of the focal task. The purpose

of the present series of experiments is to contribute further to the characterisation of this

semantic auditory distraction by exploring the boundary conditions for the occurrence of

the phenomenon.

To achieve this goal, we revisit the category-clustering paradigm in which

exemplars (e.g., "plum", "seagull") from several semantic categories (e.g., "Fruit",

"Birds") are presented for recall. Participants tend to spontaneously cluster relatively long

(e.g., 32-48 item lists), semantically categorizable lists at a greater-than-chance level at

test, even if the exemplars are randomized with respect to their categories at study (e.g.,

Bousfield, 1953; Marsh et al., 2009; Smith, Jones, & Broadbent, 1981). The advantage of

the category-clustering paradigm, over free recall of lists comprising items drawn from a

single semantic category, is that it yields a measure of semantic processing at test,

indexed by the degree to which the responses are organized by semantic category

(Bousfield, 1953; Tulving, 1968). The paradigm, therefore, provides an opportunity to

analyse how pre-existing conceptual relationships, or semantic associations, guide the

encoding and retrieval of episodic information, in contrast to (for instance) the classic

serial recall paradigm wherein the participants typically recall relatively short lists in

order of presentation and for which category-clustering would rather harm than facilitate

recall (Wetherick, 1975).

The novel characteristic of the current research is that it investigates whether: 1)

the encoding-storage stage, or the retrieval stage of the task, is the most sensitive to

semantic auditory distraction; 2) semantic auditory distraction is modulated by the

linguistic-coherence of the irrelevant material (i.e., whether it is possible to identify a

distracting 'unit' such as the individual lexical-semantic item, or whether distraction is

attributable to the processing and representation of the syntactical structure of the

irrelevant material/transitional properties between lexical-semantic items); and 3) the

output-dominance of the target and of the distracting category-exemplars, respectively,

modulates the magnitude of semantic auditory distraction.

Experiment 1

Since the 1960's, there has been a tension in the literature between whether

semantic-categorization occurs during the encoding/storage or the retrieval of categorized

lists (Allen, 1968; Slamecka, 1968). More recently, evidence in favor of the 'retrieval'

stance has accrued. For example, formal investigations of the processes underpinning

categorical-clustering have demonstrated that computationally-implemented semantic

encoding processes are not required to provide a reasonable computational fit to

category-clustering data: Semantic-categorization effects can be accounted for mainly

through semantic retrieval processes (Sirotin, Kimball, & Kahana, 2005). This implies

that the locus of semantic-categorization should be at retrieval. Indeed, semantic retrieval

processes in the absence of the requirement to encode items from the environment and store them, have nevertheless been shown to be susceptible to semantic auditory distraction: Semantic verbal fluency, a task in which participants are required to generate from long-term memory as many exemplars as possible from a specific category (e.g., "Fruit") directly from semantic memory within a time constraint, is disrupted by irrelevant speech (Hygge, Boman, & Enmerker, 2003; Jones et al., 2012; Marsh & Jones, 2010). However, perceptual manipulations during study can also influence the degree of semantic-categorization at test (Mulligan, 1999) that suggests encoding processes may play a role. In the context of semantic auditory distraction, one possibility is that meaningful irrelevant speech exerts its disruptive effect on semantic-categorization by creating a type of 'attentional blink'. In this way, selective intake of information is disrupted—perhaps impairing the shifting of attention between categories (cf. Broadbent, 1983)—giving rise to failures to organize input words into semantic categories and impairing the formulation of a retrieval plan to be executed at test (e.g., Bower, Clark, Lesgold, & Winzenz, 1969).

In Experiment 1, we investigated whether the locus of semantic auditory distraction is during encoding or retrieval by manipulating whether the irrelevant speech coincided with study/presentation or test stages of the memory task. We contrast the effect of meaningful speech with that of a meaningless sound (i.e., the speech played backwards). Reversing speech removes phonetic properties that allow lexical access, and therefore semantic processing (Sheffert, Pisoni, Fellowes, & Remez, 2002). Since the single most salient change to arise from reversing speech is the loss of meaning, finding a difference in disruption as produced by natural, forward speech in comparison with that

produced by reversed speech would imply an effect attributable to the meaningfulness of the irrelevant speech. Moreover, finding an interaction between meaningfulness and the stage of the memory task would imply that one phase of the task (e.g., study/encoding) is more susceptible to semantic auditory distraction than the other (e.g., test/retrieval) due to semantic categorization processes operating differentially at study or test. Alternatively, if equivalent disruptive effects are found regardless of whether meaningful irrelevant speech is presented at study or test, then this would imply that semantic categorization processes operate at both phases of the task. Based on what little is known (Hygge et al., 2003; Jones et al., 2012), we expected semantic auditory distraction to be similar in magnitude for sound presented at study/encoding and at test/retrieval.

Method

Participants

Forty-eight participants from Cardiff University took part in return for course credit. Each participant reported normal or corrected-to-normal vision and normal hearing and was a native English speaker. Participants were randomly assigned to one of two between-participants groups related to the timing of the exposure to irrelevant sound: Study or test.

Apparatus and Materials

To-be-remembered material. Eight instances were chosen from each of 72 categories in the Yoon et al. (2004) norms in order to construct 18 lists of 32 words, each list having 4 categories. Categories chosen had minimal category-exemplar overlap, and exemplars and categories were not repeated between or within lists in order to diminish

the effects of proactive interference (e.g., Shuell, 1968). The exemplars chosen were

sampled outside of the 10 most frequently produced instances to reduce the likelihood

that items could be recalled by simple free association or guessing (e.g., Shuell, 1969).

Categories were randomly assigned to each list but with the constraint that

associated categories (e.g., "Musical Instruments" and "Type of Music") did not appear

together. Category-exemplars within each list were arranged pseudo-randomly, so that no

two members of the same category were presented adjacently and that each category was

represented equally in each quarter of the list.

Irrelevant sound. The meaningful speech was English narrative taken from a

political essay, recorded in a female voice and sampled with a 16-bit resolution, at a

sampling rate of 44.1kHz using *Sound Forge 5* software (Sonic Inc., Madison, WI; 2000).

The recording was reversed using *Sound Forge 5* (Sonic Inc., Madison, WI; 2000) to

create meaningless speech. The speech in each of the irrelevant sound conditions was

played to participants at 65-70 dB(A) via stereo headphones that were worn throughout

the experiment.

Design

A mixed design was used with one between- and one within-participant factor.

The within-participant factor was 'Sound Condition' of which there were three levels:

forward speech, reversed speech, and quiet. The between-participants factor was 'Locus'

of which there were two levels: Study-only and Test-only. The sound conditions were

randomized as follows: The 18 lists were divided into six blocks. In each block the three

lists were randomly assigned to one of the three speech conditions. To control for order

effects, the order of irrelevant speech conditions within each block was counterbalanced

across participants such that the six possible orderings of conditions were encountered an equal number of participants within each group.

Procedure

Participants were seated at a viewing distance of approximately 60 cm from a monitor on which category-exemplars were displayed in a central position. Lists of category-exemplars appeared in lower case black 72-point *Times New Roman* font one word at a time against a white background. Each word appeared for 2 s with an inter-stimulus interval of 1 s. Retrieval was immediate with the end of the list being notified by the appearance of a red 'RECALL' cue to initiate recall.

Participants were tested in small groups of six participants. Participants were seated in individual cubicles equipped with a *Samsung Syncmaster 171S PC* and display. They were informed that they would be presented with sixteen 32-word lists, and that each list would be presented one word at a time on the computer monitor from which they were asked to memorize as many words as possible and write the words they remembered down in the order which they recalled them on recall sheets when a 'RECALL' cue appeared on the screen. Recall sheets contained 17 columns of 32 rows each. One practice trial was presented before the experimental trials. Participants were not explicitly told that the lists were categorizable. Participants were informed that they would have 2 min to retrieve as much as they could of the list and that after this time a tone would sound to signal the beginning of the next list (some 5 s later). Participants were instructed to ignore any sound that they heard through the headphones and were told that they would not be tested on its content at any point in the experiment.

Results

Recall Measures

Recall measures came in three forms: the total number of exemplars correctly recalled, the mean total number of exemplars per category correctly recalled, and the number of categories recalled (contingent on recalling one word from a category). The number of intrusions was so low as to defy statistical analysis.

Table 1 shows the results of the various recall measures in the three sound conditions. The mean scores for the total number of items correctly recalled in each condition generally show that performance was better in quiet than in both the reversed and forward speech conditions regardless of whether the speech was presented at study or test. A 3 (Sound Condition: Quiet, Reversed speech, Forward speech) × 2 (Locus: Study only, Test only) mixed ANOVA revealed a main effect of Sound Condition, $F(2, 124) = 32.12$, $MSE = 1.72$, $p < .001$, $\eta_p^2 = .34$. However, there was no main effect of Locus, $F(1, 62) = 0.47$, $MSE = 24.94$, $p = .5$, $\eta_p^2 = .007$, or any interaction between Sound Condition and Locus, $F(2, 124) = 0.05$, $MSE = 1.72$, $p = .95$, $\eta_p^2 = .001$. Post hoc tests (LSD) revealed significant differences between quiet and reversed speech ($p = .005$; $CI._{95} = .22$, 1.16), quiet and forward speech ($p < .001$; $CI._{95} = 1.35$, 2.32), and crucially, between reversed and forward speech ($p < .001$; $CI._{95} = .72$, 1.57).

The number of items per category recalled shows that performance was better in quiet than reversed speech, which in turn was better than forward speech (Table 1). A 3 (Sound Condition: Quiet, Reversed speech, Forward speech) × 2 (Locus: Study only, Test only) mixed ANOVA revealed a main effect of Sound Condition, $F(2, 124) = 23.17$, $MSE = 0.1$, $p < .001$, $\eta_p^2 = .27$. However, there was no main effect of Locus, $F(1, 62) =$

0.31, $MSE = 1.14$, $p = .58$, $\eta_p^2 = .005$, or any interaction between Sound Condition and Locus, $F(2, 124) = 0.46$, $MSE = 0.1$, $p = .63$, $\eta_p^2 = .007$. Post hoc tests (LSD) revealed significant differences between quiet and reversed speech ($p = .001$; $CI_{.95} = .08, .3$), quiet and forward speech ($p < .001$; $CI_{.95} = .26 .5$), and crucially, between reversed and forward speech ($p < .001$; $CI_{.95} = .09, .29$).

The pattern of results was different for the number of categories recalled (Table 1). Here, the means for performance in forward speech for the two locus conditions were lower than those for the reversed speech and quiet conditions, which in turn were similar to each other. A 3 (Sound Condition: Quiet, Reversed speech, Forward speech) × 2 (Locus: Study only, Test only) mixed ANOVA revealed a main effect of Sound Condition, $F(2, 124) = 10.44$, $MSE = 0.32$, $p < .001$, $\eta_p^2 = .14$. However, there was no main effect of Locus, $F(1, 62) = 1.08$, $MSE = 0.15$, $p = .3$, $\eta_p^2 = .017$, or any interaction between Sound Condition and Locus, $F(2, 124) = 0.32$, $MSE = 0.03$, $p = .73$, $\eta_p^2 = .005$. Post hoc tests (LSD) revealed significant differences between quiet and forward speech ($p < .001$; $CI_{.95} = .05 .17$), and between reversed and forward speech ($p < .001$; $CI_{.95} = .05, .17$), but not between quiet and reversed speech ($p = .61$; $CI_{.95} = -.05, .08$).

Table 1 about here

Clustering Measure

Whilst there are several potential ways of measuring semantic clustering (for a review, see Murphy, 1979), we restricted our analyses to the *Z* score (Frankel & Cole, 1971). The mathematical assumptions and processes used to compute *Z* scores are

described in more detail in Frankel and Cole's (1971) article. Essentially, this clustering measure is based on the number of runs of same-category exemplars during the output of recall. The length of a run is defined as the number of same-category items recalled in succession and single items are scored as runs of one. The number of runs is one more than the number of times the category changes during recall. Suppose *a*, *b*, *c*, and *d* represent different categories and exemplars thereof: The recalled list *aaaabbccdcbb* has six runs commencing with a run of four and terminating with a run of two. Through the process of computing the Z score, the mean and the variance of the number of runs of items within each category for each output list are computed with formulas derived from Mood (1940) and Wallis and Roberts (1957) which provide the statistics for the mean and variance in a randomly selected list of arbitrary length and categorical composition. Using this information, coupled with the number of observed runs in the output sequence, a probability of observing the number of runs (expected runs) present in a given output list can be calculated. The *Z* score can be described as: $Z = [E(L) - L]/V_L$, where *L* represents the observed number of runs and $E(L)$ represents the expected number of runs. V_L is the variance of the distribution to the predicted number of runs that depends on the number of words recalled from each category.

On the Z score measure, clustering occurs when the number of runs that are observed on the output list is significantly fewer than expected by chance. Perfect clustering e.g., *aaaabbbbccccdddd*, results in a higher Z score than imperfect clustering, e.g., *abaabccbbdcdcdda*. Positive scores indicate tendencies toward clustering. Negative *Z* scores are possible when less categorization occurs than by chance. That is, when exemplars from different categories are output adjacent to one another in the recalled list.

The Z score, therefore, has an advantage over several other methods of assessing

clustering because it enables one to tell if clustering is at an above-chance level (Frankel

& Cole, 1971).

As is typical in these types of analyses (Murphy, 1979), Z scores were computed

with all repeat and intrusion errors removed. The Z score means are lower in both sound

conditions than in the quiet condition, and lower in the forward speech compared to

reversed speech condition (Table 1). A 3 (Sound Condition: Quiet, Reversed speech,

Forward speech) × 2 (Locus: Study only, Test only) mixed ANOVA revealed a main

effect of Sound Condition, $F(2, 124) = 10.24$, $MSE = .32$, $p < .001$, $\eta_p^2 = .14$. However,

there was no main effect of Locus, $F(1, 62) = .03$, $MSE = 3.84$, $p = .87$, $\eta_p^2 = .00$, or any

interaction between Sound Condition and Locus, $F(2, 124) = .95$, $MSE = .32$, $p = .39$, η_p^2

$= .015$. Post hoc tests (LSD) revealed significant differences between quiet and forward

speech ($p < .001$; $CI_{.95} = .23\ .64$), and between reversed and forward speech ($p = .001$;

$CI_{.95} = .13, .51$), but not between quiet and reversed speech ($p = .28$; $CI_{.95} = -.09, .31$).

Discussion

The results confirm that the meaningfulness of irrelevant speech interferes with

the free recall of semantically-categorizable lists: Whilst both sound conditions decreased

the overall number of exemplars recalled, and the number of exemplars recalled by

category, meaningful irrelevant speech was more disruptive and, unlike the meaningless

sound condition, also produced disruption to the number of categories recalled and the

degree of semantic-categorization clustering demonstrated in the free recall protocols.

This aspect of Experiment 1 replicates Marsh et al. (2009) and shows that, in what

appears to be in contrast to serial short-term memory wherein the meaning of speech has little or no effect (e.g., Buchner et al., 1996; Jones & Macken, 1993), the semanticity of speech adds further disruption in a free recall setting, beyond that produced by the acoustic properties of the sound stream. This additional interference seems to specifically target semantic processing aspects in the free recall task, as the effect of meaningfulness was especially pronounced for semantic categorization of the to-be-recalled material and the number of categories recalled: two measures thought to reflect semantic or "relational" processing (Burns & Brown, 2000; Hunt & McDaniel, 1993).

Crucially, Experiment 1 demonstrates that the same degree and kind of disruption occurs regardless of whether the sound is presented at study or at test. This null effect of the locus manipulation appears to rule out an explanation based upon the idea that the meaningful irrelevant speech simply produces disruption because it causes a type of attentional blink/poor item encoding at study (cf. Broadbent, 1983). Assuming interference-free encoding at study, the impairment produced by meaningful speech at test appears to be a form of disruption related to accessibility (i.e., the ability to search for and retrieve the items from memory) rather than availability (i.e., whether the items have been encoded at study), perhaps caused by the forgetting of, or impairment in the process of generating, semantic retrieval cues (Gronlund & Shiffrin, 1986; Jones et al., 2012; Marsh & Jones, 2010).

These findings also shed light on the locus of semantic-categorization in that they suggest that it occurs at both encoding and retrieval phases of the task. Here, the participants were requested to study and recall word lists with 32 words from 4 different categories. In this context, sound locus appears to play little role to disruption:

background speech appears to be comparably as disruptive when presented at study and

when presented at test. In a previous study using lists of 16 words from a single semantic

category, Marsh et al. (2008) also found that the locus of sound presentation—at study or

test—had no consequence for the magnitude of disruption of correct recall. Hence, the

effects of background speech on correct recall appear to be robust: The fact that it is not

modulated by the locus of sound presentation is not dependent on some single detail in

the to-be-remembered material.

It is interesting to note that related speech produced more intrusions from the

irrelevant speech when presented during the presentation phase of the task in comparison

to the test phase of the task in the study by Marsh et al. (2008). Discrepancies between

the effects of background speech on correct and false recall, respectively, have been

noted in several previous studies (Marsh et al., 2008, 2009, 2012; Sörqvist et al., 2010)

and suggest that the effect on false recall is underpinned primarily by a breakdown of

source monitoring whereas the effect on correct recall may be underpinned by an

inhibitory process (Marsh et al., 2014). The results reported here add further support to

our understanding that these two effects are functionally dissimilar.

Experiment 2

Experiment 1 used coherently structured, meaningful irrelevant speech. We now

turn to investigating whether the distracting unit is the individual lexical-semantic item or

rather the syntactical processing/transitional properties between lexical-semantic items. A

comprehensive review by Underwood and Everatt (1996) proposes that unattended

information undergoes semantic analysis but only at the level of individual word

meanings. For example, in the context of dichotic listening, the latency of shadowing a word presented to the task-relevant ear was reduced when that word was preceded by a semantically-related word in the task-irrelevant ear. Thus, ignored words are semantically processed as they prime responses to target words (see also Marsh, Sörqvist, Halin, Nöstl, & Jones, 2013; Parmentier, 2008). However, whilst individual words prime semantically-related responses, there is no evidence of any priming from the meaning of sentences. For example, shadowing the sentence "put the stamp on the envelope and post the..." presented to the to-be-attended ear produces greater priming for the word "letter" than shadowing "I think we should post the...". But this pattern did not hold when the preceding context is played to the non-shadowed ear. In this case, the priming effect does not extend beyond the single related word ("post"). Therefore, Underwood and Everatt (1996) argued that there is no evidence for inter-word semantic processing of ignored speech. More specifically, they argue that listeners do not appear to recognize the deep structure of an unattended sentence or recognize the common category of words in an unattended list. This conclusion suggests that to-be-ignored random words should be at least as disruptive as coherent sentences in the case of semantic auditory distraction. This is tested in Experiment 2.

Method

Participants

Twenty participants from Cardiff University took part in return for course credit. Each participant reported normal or corrected-to-normal vision and normal hearing and was a native English speaker.

Apparatus and Materials

To-be-remembered material. Lists were constructed in an identical manner to Experiment 1.

Irrelevant sound. The text used as irrelevant speech was taken from a BBC Reith Lecture (Brave Old World by Tom Kirkwood, 2001). From this text, two order-of-approximations to English were created by removing every sixth (third order-approximation) and second (fifth-order approximation) word in the text passage and randomly reinserting them into the spaces left in the text. The words used to construct each order of approximation to English (as well as the natural text) were recorded by a male speaker in list intonation (Tun, O'Kane, & Wingfield, 2002) at a rate of approximately 3 words per sec. This was presented during the study phase of the task.

Design

A single-factor within-participants design was used. The within-participant factor was 'Sound Condition' of which there were four levels: Text-speech, 3rd-, and 5th-order of approximation to English speech and a quiet condition. The irrelevant speech was played during both study and test phases of the task. Other aspects of the design, and the procedure, were identical to Experiment 1.

<div align="center">Results</div>

Recall Measures

Recall measures were the same as for Experiment 1. In general, the mean scores for the total number of category-exemplars recalled correctly were greater in the quiet condition than in the sound conditions (Table 2). An ANOVA confirmed a main effect of Sound Condition on the total number of category-exemplars correctly recalled, $F(3, 57) =$

10.16, $MSE = 3.70$, $p < .001$, $\eta_p^2 = .35$. Post hoc tests (LSD) revealed significant

differences between quiet and text ($p < .001$, $CI_{.95} = 1.23, 3.34$), quiet and 3rd-order

approximation ($p = .002$, $CI_{.95} = 1.33, 4.85$), and between quiet and 5th-order

approximation ($p = .001$, $CI_{.95} = 1.19, 4.04$). There was no significant difference between

text and 3rd-order approximation ($p = .15$, $CI_{.95} = -.32, 1.92$), or between text and 5th-

order approximation ($p = .52$, $CI_{.95} = -.71, 1.36$), and 3rd-order approximation did not

differ from 5th-order approximation ($p = .37$; $CI_{.95} = -.61, 1.56$).

An identical pattern emerged for the number of exemplars per category recalled,

$F(3, 57) = 5.12$, $MSE = .21$, $p = .003$, $\eta_p^2 = .21$. Post hoc tests (LSD) demonstrated that

the mean scores in quiet were significantly different from those in text ($p = .011$, $CI_{.95} =$

$.15, .54$), 3rd-order approximation ($p = .018$, $CI_{.95} = .10, .93$), and 5th-order approximation

($p = .005$, $CI_{.95} = .15, .75$). Text did not differ from 3rd-order approximation ($p = .29$;

$CI_{.95} = -16, .51$), or 5th-order approximation ($p = .33$; $CI_{.95} = -.12, .33$), and 3rd-order

approximation was not significantly different from 5th-order approximation ($p = .63$; $CI_{.95}$

$= -.22, .35$).

Analysis of the number of categories recalled again followed the same pattern,

$F(3, 57) = 11.6$, $MSE = .035$, $p = .003$, $\eta_p^2 = .38$. Post hoc tests (LSD) demonstrated that

the mean scores in quiet were significantly different from those in text ($p = .004$, $CI_{.95} =$

$.09, .39$), 3rd-order approximation ($p = .002$, $CI_{.95} = .10, .38$), and 5th-order approximation

($p = .005$, $CI_{.95} = .07, .35$). Text did not differ from 3rd-order approximation ($p = 1.0$;

$CI_{.95} = -19, .19$), or 5th-order approximation ($p = .63$; $CI_{.95} = -.13, .08$), and 3rd-order

approximation was not significantly different from 5th-order approximation ($p = .76$; $CI_{.95}$

$= -.14, .19$).

Clustering Measure

The *Z* score means were generally higher in quiet condition than in the sound conditions. An ANOVA confirmed a main effect of Sound Condition on *Z* scores, $F(3, 57) = 9.42$, *MSE* = .44, $p < .001$, $\eta_p^2 = .33$. Follow-up post hoc tests (LSD) revealed significant differences between quiet and text ($p < .001$; *CI*.95 = .39, 1.15), quiet and 3rd-order approximation ($p < .001$, *CI*.95 = .60, 1.46), and quiet and 5th-order approximation ($p = .001$, *CI*.95 = .38, 1.34). Text was not significantly different from 3rd-order approximation ($p = .18$, -.13, .65) or 5th-order approximation ($p = .72$, *CI*.95 = -.43, .62), 3rd-order approximation was not significantly different from 5th-order approximation ($p = .42$, *CI*.95 = -.26, .59).

Table 2 about here

Discussion

The results of Experiment 2 demonstrate that the effects attributable to the meaningfulness of irrelevant speech observed in Experiment 1 are not due to the highly structured nature of English narrative speech, nor are they attributable to the presence of information such as strong prosodic contours and co-articulatory cues within the speech signal: Simply conveying semantic properties at the lexical level was enough to significantly disrupt free recall. Therefore, the distracting unit responsible for semantic

auditory distraction appears to be the individual word rather than a more global level of processing.

Attention to the sound is an unnecessary condition for producing disruption to serial recall (Macken, Tremblay, Houghton, Nicholls, & Jones, 2003). The results of Experiment 2 suggest that attention to the background speech is, similarly, not a pre-requisite for disruption to free recall of semantic information. Consistent with the finding that reading comprehension is as disrupted by a speech stream of random words than it is by a speech stream that consists of coherent prose (Martin et al., 1988), sentences and individual words do not differ in their disruptive effects on free recall, even though a more coherent speech (e.g., sentences) would presumably be more likely to capture, or voluntarily engage, a participant's attention (Tun et al., 2002). Naturally, the fact that no difference was found between sentences and individual words may be due to weak statistical power. However, the absence of a difference between these two conditions is found systematically across studies (cf. Martin et al., 1988). Furthermore, any potential difference between these two background conditions is clearly not as easily detected as other robust findings in the context of semantic auditory distraction, making the explanatory compass of the attention capture view limited, at best. Taken together, it seems as if meaningful irrelevant speech produces disruption to free recall via some means other than diverting the locus-of-attention away from the target memoranda.

One possibility is that speech interferes with the reproduction of the visually-presented material via the automatic, obligatory processing of irrelevant speech meaning (Jones et al., 2012; Marsh & Jones, 2010). The speech material activates representations in semantic memory that compete with the output (e.g., organization of responses for

retrieval: output planning) of the to-be-recalled items (Marsh et al., 2008, 2009; Sörqvist et al., 2010). According to this interference-by-process view, the semantic similarity between the to-be-recalled and the to-be-ignored items should modulate the degree of disruption. Moreover, speech that comprises high output dominance exemplars should be particularly disruptive when the to-be-recalled items comprise low output dominance exemplars because they activate highly primed responses that must be inhibited for successful retrieval of the (low dominance) targets (Anderson, 2003; Marsh et al., 2013). For instance, in the case when the to-be-ignored and the to-be-remembered material share category membership, irrelevant exemplars (e.g., 'robin,' 'sparrow') are candidates for retrieval because they convey information that yields congruent-yet-incongruent response information ("robin" is a "Bird" but was not presented on the to-be-recalled list). Successful retrieval of the to-be-remembered items thus requires inhibition of the irrelevant exemplars. An effect attributable specifically to between-stream semantic *similarity* could thus be explained as the result of functional, selective attention processes that act to inhibit highly compatible responses from being retrieved (see also Marsh et al., 2008, 2009, 2012).

<div style="text-align:center">Experiment 3</div>

In Experiment 3, we again manipulate semantic aspects of the material, but this time by presenting category-exemplars as to-be-remembered and to-be-ignored material. We varied the semantic-similarity between the to-be-remembered and the to-be-ignored material and the output dominance of the category-exemplars it comprises. Output-dominance reflects the relative strength with which an item (high: "orange", low: "guava") represents the category that it is a member of ("Fruit"), as indexed by the

number of people who list it in norming studies (Van Overschelde, Rawson, & Dunlosky, 2004). Prior research has demonstrated that clustering—an index of semantic processing—is reduced for low-dominance as opposed to high-dominance category-exemplars (Bousfield & Puff, 1964; Cofer, Bruce, & Reicher, 1966; Hudson, 1969). Low output-dominance to-be-remembered material should, thus, be associated with less semantic processing.

In Experiment 3, the output dominance of the to-be-remembered material, in relation to the to-be-ignored material, was manipulated such that: a) the to-be-remembered and to-be-ignored exemplars were both high output dominant (HH); b) the to-be-remembered material was high output dominant, but the to-be-ignored material was low output dominant (HL); c) the to-be-remembered material was low output dominant and the to-be-ignored material was high output dominant (LH); and d) both the to-be-remembered and the to-be-ignored material were low output dominant (LL). A crucial additional manipulation was that the to-be-ignored material could either belong to the same semantic category (related condition) as one of the to-be-remembered categories or to a different category (unrelated condition). This manipulation made it possible to examine the role of competition.

Method

Participants

Eighty undergraduate students from Cardiff University participated in exchange for course credit. All participants reported normal or corrected-to-normal vision and normal hearing.

Apparatus and Materials

To-be-remembered material and irrelevant sound. The to-be-recalled exemplars consisted of 8 lists of 32 words with each list comprising 8 words from 4 different categories (Van Overschelde et al., 2004). Category-exemplars for the high output-dominant to-be-recalled set came from even ranked positions (2 to 16) in the Van Overschelde et al. (2004) category norm lists (e.g., for the category "fruit," (one of four categories on the to-be-remembered list) the category-exemplars chosen would be orange [2], grape [4], peach [6], etc.). For the low output-dominant to-be-recalled set, category-exemplars came from positions 18 to 32 (e.g., for the category "fruit", the to-be-remembered lists would comprise (cantaloupe [18], lime [20], papaya [22], etc.).

Irrelevant items chosen as high output-dominant were taken from the odd ranked positions (1 to 15) in the Van Overschelde et al. (2004) category norm lists [e.g., for the category "fruit," the category-exemplars chosen would be apple [1], banana [3], pear [5], etc.]. Low dominant items were chosen from positions 16 to 32 [e.g., for the category "fruit", the to-be-remembered lists would comprise blueberry [17], raspberry [19], tangerine [21], etc.).

Exemplars in the to-be-ignored sequence were either related to one of the categories represented in the to-be-recalled list or unrelated to all of the categories represented in the to-be-recalled list. All to-be-ignored exemplars were presented at the rate of one word 750 ms, with their sequence looped to create a total length that lasted from the start to the end of the trial and they were only played during Study of the to-be-recalled exemplars. The to-be-ignored exemplars were arranged pseudo-randomly so that each exemplar was presented once every eight exemplars and no category exemplar was

followed by itself. For each participant, none of the to-be-ignored exemplars were used as to-be-recalled exemplars.

Design

A mixed design was adopted with Relatedness (related and unrelated) as the within-participant factor and Dominance [high dominant to-be-recalled and high dominant to-be-ignored (HH), high dominant to-be-recalled and low dominant to-be-ignored (HL), low dominant to-be-recalled and high dominant to-be-ignored (LH) and low dominant to-be-recalled and low dominant to-be-ignored (LL)] as the between-participants factor. For each group, five possible to-be-recalled presentation orders were created with four participants assigned to each. This procedure was performed twice, once for the related and once for the unrelated condition. Finally, half of all participants received the high, and half received the low, to-be-ignored exemplars first.

Procedure

This was the same as Experiments 1 and 2.

Results

Recall Measures

Intrusion and repetition rates were too low to be subject to statistical analysis. Table 3 illustrates the results of the various recall measures in the two sound conditions and the four dominance conditions. The mean scores for the total number of items correctly recalled shows that recall for high output-dominant exemplars was better than for low output-dominant exemplars. Further, performance was also better when ignoring

unrelated, compared with related, exemplars. However, consistent with our predictions, this relatedness effect was larger when the ignored exemplars were high output-dominant.

A 2 (Relatedness) × 4 (Dominance) mixed ANOVA with category-exemplars correctly recalled as dependent measure revealed a significant main effects of Relatedness, $F(1, 76) = 37.05$, $MSE = 2.57$, $p < .001$, $\eta_p^2 = .33$, and Dominance, $F(3, 76) = 26.28$, $MSE = 12.03$, $p < .001$, $\eta_p^2 = .51$, as well as a significant Relatedness × Dominance interaction, $F(3, 76) = 8.13$, $MSE = 2.57$, $p < .001$, $\eta_p^2 = .24$. To tease the interaction apart, simple effects analyses (LSD) were first conducted on differences between related and unrelated speech. Significant differences were found for the HH ($p < .001$, $CI_{.95} = .97, 2.99$) and LH conditions ($p < .001$, $CI_{.95} = 2.34, 4.36$), but not for HL ($p = .11$, $CI_{.95} = -.19, 1.84$) or LL ($p = .96$, $CI_{.95} = -.99, 1.04$) conditions. As predicted, the relatedness effect was only evident when the ignored exemplars were high output-dominant as they provided good representations of their super-ordinate category which required inhibition to prevent them from being erroneously selected for recall. By contrast, ignoring low output-dominant exemplars did not provide a strong representation of their super-ordinate category and consequently they did not require inhibition.

To further examine the significant interaction between Dominance and Relatedness, simple effects analyses (LSD) were conducted on differences between dominance conditions. For the related condition, significant differences were found between HH and LH ($p < .001$, $CI_{.95} = 5.27, 8.48$), HH and LL ($p < .001$, $CI_{.95} = 1.84, 5.06$), HL and LH ($p < .001$, $CI_{.95} = 4.43, 7.65$), HL and LL ($p = .002$, $CI_{.95} = 1.004, 4.22$), and LH and LL ($p < .001$, $CI_{.95} = -5.03, -1.82$), but not HH and HL ($p = .3$, $CI_{.95} = -.771, 2.45$). For the unrelated condition, significant differences were found between HH

and HL (p = .03, $CI_{.95}$ = .197, 3.78), HH and LH (p < .001, $CI_{.95}$ = 3.71, 7.29), HH and

LL (p < .001, $CI_{.95}$ = 3.61, 7.19), HL and LH (p < .001, $CI_{.95}$ = 1.72, 5.30), HL and LL

(p < .001, $CI_{.95}$ = 1.622, 5.20), but not between LH and LL (p = .91, $CI_{.95}$ = -1.89, 1.69).

This demonstrates that recall was generally better for when the to-be-recalled material

was high output-dominant. This supports the notion that the high output-dominant

category-exemplars received more semantic/elaborative encoding than the low output-

dominant items (Bousfield & Puff, 1964; Cofer et al., 1966; Hudson, 1969).

To assess whether the relatedness effect of irrelevant speech was specific to the

category that was ignored in the related condition or whether the apparent retrieval

inhibition spread to the other category exemplars in the to-be-recalled list, a further 2

(Category Match: Matching versus Non-Matching) × 4 (Dominance) mixed ANOVA

was conducted. There was no main effect of Match, $F(1, 76)$ = .86, MSE = .26, p = .36,

η_p^2 = .01, but of Dominance, $F(1, 76)$ = 25.33, MSE = .88, p < .001, η_p^2 = .5, whereas

there was no interaction between the factors, $F(3, 76)$ = .78, MSE = .33, p = .51, η_p^2 =

.03. This lack of a category match effect suggests that the retrieval impairment spreads to

other exemplars that, although they are unrelated in content to the to-be-ignored

exemplars, are being processed similarly; that is, by using semantically categorizable

information (see also Marsh et al., 2009).

The number of items per category correctly recalled generally mirrored the mean

scores for the total number of items correctly recalled (Table 3). More items per category

were recalled when the to-be-recalled items were high output-dominant than low output-

dominant and the related speech appeared to produce most disruption when the to-be-

ignored material was high output-dominant. The ANOVA revealed an identical pattern

for the number of exemplars per category recalled. There was a main effect of

Relatedness, $F(1, 76) = 41.36$, $MSE = .15$, $p < .001$, $\eta_p^2 = .35$, and of Dominance, $F(1,$

$76) = 16.04$, $MSE = .884$, $p < .001$, $\eta_p^2 = .39$, and the Relatedness $\times$ Dominance

interaction was also significant, $F(3, 76) = 10.37$, $MSE = .15$, $p < .001$, $\eta_p^2 = .29$. Simple

effects analysis revealed a relatedness effect for HH ($p < .001$, -.738, -.252), LH ($p <$

.001, $CI_{.95} = -.643, 1.13$) and a marginally significant effect for HL ($p = .079$, $CI_{.95} = -$

.459, .026). There was clearly no relatedness effect for LL ($p = .81$, -.214, .272).

Additionally, simple effects analysis (LSD) for the related condition revealed significant

differences between HH and LH ($p < .001$, $CI_{.95} = 1.08, 1.96$), HH and LL ($p = .001$,

$CI_{.95} = .362, 1.25$), HL and LH ($p < .001$, $CI_{.95} = .74, 1.63$), HL and LL ($p = .038$, $CI_{.95} =$

.027, .911), and LH and LL ($p = .002$, $CI_{.95} = -1.16, -.27$), but not HH and HL ($p = .14$,

$CI_{.95} = -.107, .778$). For the unrelated condition, significant differences were found

between HH and HL ($p = .01$, $CI_{.95} = .151, 1.08$), HH and LH ($p < .001$, $CI_{.95} = .67$,

1.59), HH and LL ($p < .001$, $CI_{.95} = .87, 1.79$), HL and HH ($p = .03$, $CI_{.95} = .052, .977$),

and HL and LL ($p = .003$, $CI_{.95} = .252, 1.177$), but not LH and HH ($p = .39$, $CI_{.95} = -.263$,

.662).

The number of categories recalled for each condition is also found in Table 3.

Generally, categories were recalled more poorly only when the to-be-recalled information

was low output-dominant and the to-be-ignored information was high output-dominant.

Related speech did not appear to have an additional effect on category recall. Analysis of

the number of categories recalled failed to reveal a significant effect of Relatedness, $F(1,$

$76) = .56$, $MSE = .045$, $p = .46$, $\eta_p^2 = .007$. However, there was a main effect of

Dominance, $F(3, 76) = 49.51$, $MSE = .021$, $p < .001$, $\eta_p^2 = .66$, but the interaction

between these variables was not significant, $F(3, 76) = .023$, $MSE = .045$, $p = .995$, $\eta_p^2 = .001$. Generally, categories were more poorly recalled in the LH group, which suggests that the disruption produced by semantic relatedness to category recall only occurs when the to-be-recalled material is relatively low in output-dominance (and thereby more difficult to categorize) and the to-be-ignored material is high in output-dominance (and therefore more likely to elicit automatic semantic activation and processing).

Clustering Measure

Clustering scores were higher for high output-dominant items (Table 3). Related speech seemed to disrupt clustering only when the to-be-recalled material was low in output-dominance and the irrelevant items were high in output-dominance, or when both to-be-recalled and irrelevant items were high in output-dominance. An ANOVA confirmed a main effect of Relatedness, $F(1, 76) = 14.48$, $MSE = .49$, $p < .001$, $\eta_p^2 = .16$, and a main effect of Dominance, $F(1, 76) = 19.39$, $MSE = 2.3$, $p < .001$, $\eta_p^2 = .43$. The Relatedness × Dominance interaction was also significant, $F(3, 76) = 7.4$, $MSE = .49$, $p < .001$, $\eta_p^2 = .23$.

A simple effects analysis (LSD) revealed a marginally significant relatedness effect for HH ($p = .051$, $CI._{95} = -.876, .001$), a significant relatedness effect for LH ($p < .001$, $CI._{95} = -1.688, -.811$), but not for HL ($p = .6$, $CI._{95} = -.324, .553$) or LL ($p = .64$, $CI._{95} = -.541, .336$). A simple effects analysis for the related condition also revealed significant differences in clustering scores between HH and HL ($p = .003$, $CI._{95} = .372, 1.702$), HH and LH ($p < .001$, $CI._{95} = 2.048, 3.378$), HH and LL ($p < .001$, $CI._{95} = 1.321, 2.650$), HL and LH ($p < .001$, $CI._{95} = 1.011, 2.340$), HL and LL ($p = .006$, $CI._{95} = .283, 1.613$), and between LH and LL ($p = .032$, $CI._{95} = -1.392, -.062$). A simple effects

analysis for the unrelated condition revealed significant differences between HH and HL ($p < .001$, $CI._{95} = .775, 2.404$), HH and LH ($p < .001$, $CI._{95} = 1.085, 2.714$), HH and LL ($p < .001$, $CI._{95} = 1.505, 3.134$), a marginally significant difference between HL and LL ($p = .078$, $CI._{95} = -.084, 1.545$), but not between HL and LH ($p = .45$, $CI._{95} = -.504, 1.125$) or between LH and LL ($p = .31$, $CI._{95} = -.395, 1.234$).

Table 3 about here

Discussion

Experiment 3 contributes to the growing number of findings that have demonstrated between-stream semantic similarity effects in cross-modal auditory distraction (Beaman, 2004; Bell et al., 2008; Marsh et al., 2008, 2009, 2012; Neely & LeCompte, 1999; Sörqvist et al., 2010). Crucially, the disruption to clustering produced by related speech occurred only when the irrelevant items were strong competitors for recall (i.e., high output dominant category members), but not when the to-be-recalled and irrelevant items were unlikely to compete (i.e., when the to-be-recalled items were high and the to-be-ignored items were low output dominant). This indicates that a by-product of competition is reduced clustering and is consistent with the notion of a strategy disruption mechanism (Marsh et al., 2009).

General Discussion

The results of Experiments 1-3 add to a small body of existing evidence demonstrating that to-be-ignored sound is semantically processed, even though attention

is allocated to the visual modality, and this involuntary semantic analysis disrupts the focal task when incompatible with the task requirements (e.g., Parmentier, 2008). As opposed to the effects of sound on serial short-term memory (e.g., Buchner et al., 1996; Marsh et al., 2008; Sörqvist et al., 2010), the meaningfulness of speech is endowed with disruptive power, that goes beyond the sound's acoustic properties, in the context of free recall of visually-presented category exemplars. This semantic auditory distraction occurs regardless of whether the irrelevant sound is presented during study or test phases of the task (Experiment 1), occurs at the lexical, rather than inter-lexical, level of the speech (Experiment 2), is greater in magnitude when the speech conveys information that is semantically-related to the to-be-recalled material, and is particularly pronounced when the to-be-ignored speech words are output dominant category members (Experiment 3). Furthermore, the semantic properties of the speech stream adds to the impairment of the total number of exemplars recalled, the number of exemplars per category recalled and the number of categories recalled, but the disruptive effect of similarity between the to-be-remembered and to-be-ignored material does not seem to influence the number of categories recalled.

A note of caution of required, however, as the effects of background speech on categories recalled was very small in magnitude, at least in view of the magnitude of the difference between means. Looking at the data for Experiment 1, just over 1 participant in 2 (56%) recalled fewer categories in the meaningful speech condition compared to quiet. Similarly, for Experiment 2, 1 participant in 2 (50% of participants) recalled fewer categories in the text condition as compared to the quiet condition. The figure approximated 2 participants in 3 (70%) for the 3rd order approximation to English

condition as compared with quiet, whilst in the 5th order approximation to English, the

figure was just over 1 in 2 participants (55%). Hence, although the differences between

means were rather small (see Tables 1-3), the difference between conditions was in the

same direction for the majority of participants. Moreover, the effect was found in three

independent experiments. We can conclude that the effects seen here are small, but both

reliable and replicable.

The locus of distraction

The current study speaks to the mechanisms of semantic auditory distraction and

the nature of irrelevant sound processing. Random words, which together constituted a

speech stream with low approximation to English, caused as much disruption as a

coherent speech stream with high approximation to English (Experiment 2). This result

accords well with those reported by Martin et al. (1988) who found that an irrelevant

passage comprising a coherent story produced no more disruption of reading

comprehension than randomly arranged words that were taken from the same story. This

suggests that neither the continuity of speech (i.e., the presence of co-articulatory cues)

nor its inter-word semantic properties (i.e., semantic transitional probabilities) govern the

disruption it produces to semantically-based focal tasks. Rather, meaningful irrelevant

speech produces disruption attributable to its lexical item properties, not its supra-lexical

properties.

Of course, even though these results suggest that the disruption is likely to be at

the level of individual word meanings (e.g., the recognition of the meaning of to-be-

ignored words), the consequence of this conflict can be observed at a supra-word level;

that is, manifest in the poorer semantic-integration (or clustering) of to-be-recalled words

as shown in Experiments 1-3. Indeed, any impairment of the processing of the meaning of words will affect integration or clustering because the processes underpinning semantic-organization requires the recognition of the associations between incoming stimuli and existing knowledge structures (Underwood & Everatt, 1996). More specifically, meaningful irrelevant speech interferes with the focal task processing at the level of the individual meanings of to-be-recalled items [or the use of the meanings of those items for, say, semantic organization and the establishment at encoding, or use at retrieval, of higher-order semantic encodings that can be used as a retrieval plans to enable inter-category transitions (e.g., Bower et al., 1969)].

The comparison between effects of sound at encoding/study and retrieval/test provides further support for the assumption that speech produces its disruption by corrupting the semantic organization at an output level. As shown here, speech appears to be as disruptive to semantic-categorization (as well as correct recall of exemplars and categories) when presented at the encoding phase of the short-term memory task as when presented at retrieval (Experiment 1). Thus, output planning processes operating at the encoding as well as at the retrieval phase are disrupted in the presence of background speech. This is entirely consistent with the notion that background speech impairs generation of category exemplars from long-term memory (e.g., long-term semantic memory; Jones et al., 2012) and with recent findings demonstrating that speech interferes with random generation when presented concurrently with the execution of the task, but only when the speech comprises items that prime dominant responses which then have to be overruled to achieve successful random generation (Marsh et al., 2013).

If the disruption of focal semantic processes is a by-product of the semantic processing of irrelevant speech and this is intrinsically related to how irrelevant speech is processed, can the mechanism of interference be specified? Underwood and Everatt (1996) sketch out some ways in which unattended speech processing can influence focal task processing. According to the non-selective access hypothesis (Underwood, 1981), unattended words gain access to the lexicon regardless of the relationship with the target stimulus but the effect of between-sequence semantic similarity arises during processes after recognition. Candidate affected processes include selection of the recognized lexical token, and selection of the response. More specifically, when two stimuli (a target word and an unattended word) activate their lexical representations, one of them must be selected as the basis for the organization of the response. Between-sequence semantic dissimilarity does not have a robust effect on the selection of an appropriate lexical token in comparison to between-sequence similarity (Experiment 3), because the two sources of activation in the lexicon point to semantically distant words and thus their separation poses little processing difficulty. However, when the unattended and attended material is associated, then the selection of the target—or its separation from activated unrelated items—may be impeded by the activity caused by its near neighbour. According to Underwood and Everatt (1996), different effects of unattended items that are unrelated or related to those to-be-recalled items are produced due to effects on different stages of processing: The disruption produced by between-sequence semantic similarity, for example, could be thought to have its effect at a processing stage whereby lexical entries are accessed (via automatic spreading activation), whereas unrelated words may affect a processing stage whereby a target has to be selected in preparation for a response. One

extension of Underwood and Everatt's (1996) activation-based theory would be to

assume that competition for retrieval arises to the extent that an item receives automatic

spreading activation from list responses, thereby suggesting retrieval orders that impact

on the organization and semantic clustering of responses (cf. Marsh et al., 2009).

Implications for theories of semantic auditory distraction

      The findings reported here are difficult to reconcile with the view that speech

disrupts short-term memory of semantic information because it captures attention (Bell,

Röer, Dentale, & Buchner, 2012; Buchner Mehl, Rothermund, & Wentura, 2006). On this

account (Cowan, 1995), acoustic changes-in-state from one irrelevant item to the next

causes an orienting response (OR) away from the focal task regardless of the processes

the focal task engages. There is already extensive evidence against this account in the

context of visual-verbal serial recall (see Hughes, Vachon, & Jones, 2005). In the context

of semantic auditory distraction, the attentional capture account could suppose that

attentional capture is necessary pre-requisite for semantic processing of the distractors

and thereafter to yield semantic interference (e.g., Parmentier, 2008). Under the umbrella

of this assumption, one must propose that distractors that are semantically related to

memoranda may be more easily activated and therefore likely to capture attention

because of their semantic associations to the attended material. However, we feel that this

account cannot easily explain why between-stream semantic similarity only exerts its

disruptive effect when semantic-categorization is emphasised as a recall strategy (Marsh

et al., 2009). It seems logical that the attentional capture account would predict an effect

of between-sequence semantic similarity regardless of task instruction because between-

sequence semantic similarity on this account is also thought to give rise to ORs [based

upon priming of semantic features, cf. Cowan, 1995; such priming is thought to be

underpinned by automatic semantic priming (e.g., J. H. Neely, 1976) that occurs 'full

blown' regardless of the focal task (see J. H. Neely & Kahan, 2001)]. Moreover, on the

attentional capture account, it is not readily obvious why text speech does not produce

any more disruption than an approximation to English (Experiment 2).

Moreover, extant research has shown that sound, specifically designed to capture

attention, such as the "m" in the sequence "kkkkkkkmkk", only produces disruption when

presented during the encoding phase, not when presented during a retention interval

(Hughes et al., 2005, 2007). The mechanism underpinning disruption by natural speech—

as opposed to abruptly irregular speech—appears to be insensitive to this type of

manipulation. For instance, both serial recall (e.g., Jones & Macken, 1993) and free recall

(e.g., Beaman, 2004; Sörqvist et al., 2010) are impaired by speech presented during a

retention interval between encoding and retrieval, as well as when presented at

study/encoding and retrieval/test (Experiment 1). Because of these findings, we argue

that natural speech (without any apparent pop-out element) does not disrupt free recall of

semantic information due to an attentional blink that prevents successful encoding

(Broadbent, 1983) or because it captures attention from the focal task (Bell et al., 2012).

One alternative explanation of the finding that natural speech produced no more

distraction than random speech in Experiment 2 is that the focal task of memorising the

categorizable list does not tap syntactical processing, and thereby the syntactical

properties of irrelevant speech are rendered impotent in their capacity to impair focal task

processing. However, this suggestion is at odds with the findings of Martin et al. (1988)

that show that the presence of syntax within the irrelevant speech does not impair reading

comprehension which would obviously require syntactical processing. In order to account for the finding that syntax is unimportant as a determinant of distraction, we suppose that unattended information undergoes semantic analysis only at the level of individual word meanings. However, in this respect the attentional capture account could also be defensible by making the similar assumption that the syntactical properties of the irrelevant speech are not processed when the speech is ignored, thereby explaining why they are unable to determine attention capture. At odds with the latter argument, however, is the finding that syntactical violations within auditory sentences elicit early left anterior negativity in event-related potential studies, regardless of whether the speech is attended or to-be-ignored, which suggests that syntactical processing occurs in a manner invariant of selective attention (Maidhof & Koelsch, 2011). In light of this, perhaps the most logical assumption is that interference arises between the semantic processing of individual word meanings because these a processed *prior* to the processing of syntax (cf. Zhang, Li, Piaoa, Liud, Huanga, & Shu, 2013) and thereafter any processing of irrelevant syntax will be weakened or rendered redundant in its capacity to disrupt focal task processing, even when the task requires syntactical processing (e.g., reading comprehension; Martin et al., 1988).

Generally, attentional resource accounts also posit that disruption of a primary task is due to the sound capturing attention away from the memory task. In essence, these accounts assume that the presence of irrelevant speech creates a dual task setting in which ignoring speech acts as a secondary task, drawing away 'resources' from the memory task (Neath, 2000). On the face of it, the results obtained here resemble those reported with divided attention studies that have also used categorizable word lists (Baddeley,

Eldridge, Lewis, & Thomson, 1984; Moscovitch, 1994). However, it would be premature to conclude that the semantic content was simply creating a divided attention setting because meaningful, as compared with meaningless, speech does not differentially disrupt serial recall of digits (Jones et al., 1990), or the recall of categorizable lists participants are instructed to recall in serial order (Marsh et al., 2008, 2009). As is clear in the case of other tasks (Marsh et al., 2013; Marsh, Sörqvist, Beaman, & Jones, 2013), these attentional resource-based accounts do not explain why the nature of distraction effects are intricately related to the nature of the focal task activity.

The findings reported here are also difficult to reconcile with an interference-by-content approach to auditory distraction (Gathercole & Baddeley, 1993; Neath, 2000). According to this view, the similarity in content between the speech items and the to-be-recalled items (e.g., whether the two sound alike) is the mechanism that determines the magnitude of disruption, not the cognitive processes that operate on those items. One finding that is especially troublesome for the interference-by-content approach is that the between-sequence semantic similarity effect found in Experiment 3 did not produce category-specific impairment (see also, Marsh et al., 2009). That disruption produced by the irrelevant exemplars that matched one of the four categories represented by the to-be-remembered list was not restricted to that matching category. This flies in the face of the category-specific impairment that would be expected if disruption was a passive side effect of the co-existence of to-be-remembered and to-be-ignored items within the same representational space (e.g., Gathercole & Baddeley, 1993; Neath, 2000).

Another finding that deserves further exposition is that low output-dominant distractors produced more disruption than high output-dominant distractors when they

were unrelated to the to-be-recalled exemplars (see Experiment 3). Assuming that output-dominance is positively related to word frequency in the English language (e.g., Mervis, Catlin, & Rosch, 1976), we suggest that this finding may be related to the distractor-frequency effect in the domain of picture-word interference (Miozzo & Caramazza, 2003), whereby the naming of a picture is slowed to a greater extent by the presence of a low- as compared to high-frequency distractor-word that is either embedded within the picture (intra-modal distraction) or presented auditorily (cross-modal distraction; Hantsch, Jescheniak, & Schriefers, 2009). According to the response-exclusion account, high frequency distractors gain access to an output buffer very quickly and take less time to remove whereas low frequency distractors take a long time to occupy the response buffer and must be removed before the target word (picture name) can be produced. Our proposal here is that, in the case of recall of word lists, the speech production system is co-opted to support recall through supporting covert articulatory rehearsal (e.g., Tan & Ward, 2000). Speech production processes may play a role in supporting the selection process ensuring that the appropriate candidate item is selected from an activated conceptual representation (e.g., the semantic category from which the to-be-recalled items are taken). Our notion here is that the greater demand that the focal task has on the speech-production system, the more likely it is that distractor-variables--such as distractor-frequency--that impair performance on tasks that require speech-production will assume their disruptive potency (see also, Beaman, Marsh, & Jones, 2014; Marsh, Crawford, Pilgrim, Sörqvist, & Hughes, 2014).

Although speculative, we suggest that, on the rare occasions in which lexical factors associated with irrelevant speech influence visual-verbal serial recall, they do so

because semantically rich material, such as words, are presented for recall thereby recruiting the speech production system. For example, Buchner and Erdfelder (2005) presented verbal word stimuli as targets and coupled this with the requirement for articulatory output (vocal responses). They found that low frequency distractors impaired serial recall more than high frequency distractors. Arguably, the nature of the target stimuli used by Buchner and Erdfelder, combined with the output mode, could have recruited the speech production system (e.g., to select the appropriate item from long-term memory for covert rehearsal, or to prepare targets for spoken output; but see Elliott & Briganti, 2012). We note that this account of the distractor-frequency effect within serial recall is at odds with the account offered by Buchner and Erdfelder (2005) whom assume that low frequency distractors impair performance because they draw limited attentional resources away from the focal task (Buchner & Erdfelder, 2005; cf. Beaman et al., 2014; but see Elliott & Briganti, 2012).

Although more detailed accounts of semantic auditory distraction are clearly required, as a whole, the results reported here appear consistent within the process-oriented, strategy-based account of auditory distraction (Hughes & Jones, 2005; Jones & Tremblay, 2000; Marsh et al., 2008, 2009, 2012; Neumann, 1996). The process-oriented view holds that the automatic, obligatory processing of serial order of changing acoustic elements of sound conflicts with the deliberate processing of serial order in the primary task, thus impairing visual-verbal serial recall. Here, the *interference-by-process* view suggests that the interference may be construed as the residual cost of mechanisms (perhaps inhibition; see Hughes & Jones, 2003; Marsh et al., 2012) which are designed to resolve a 'competition-for-action' conferred by the order cues generated from irrelevant

and relevant sources of information. This approach neatly explains why neither the meaningfulness of irrelevant speech nor between-stream semantic similarity plays a role in the disruption when serial recall is instructed: In this case, it is the information that the irrelevant sound yields about serial order, not its meaning, that is broadly compatible with the action (or process) of serial rehearsal. The interference-by-process approach also explains why the meaning of speech becomes disruptive to the performance of free recall tasks only when semantic-categorization is an obvious or instructed strategy. When the primary task involves semantic encoding and retrieval processes—unlike the case with serial recall—the irrelevant semantic information extracted from the speech—here shown to operate at a lexical-item level—produces competition for these processes. Impairment can thus be understood in terms of a relative difficulty in selecting the correct source of semantic information as they both compete for the category of action being called for in the semantic recall task.

That between-stream semantic similarity produces more interference than mere meaningfulness (Experiment 3) is particularly supportive of a process-oriented account. From an inhibitory point of view, irrelevant speech that is semantically-similar to to-be-recalled material, specifies highly compatible, but ultimately response-inappropriate, information in the context of the semantic recall task and is thus likely to require greater inhibition. However, greater inhibitory demands are only required when the irrelevant speech items are high in terms of output-dominance, with the inhibitory process serving the function of supporting the retrieval of to-be-recalled items: Speech that comprises high output dominance exemplars appears to prime responses that have to be overruled,

possibly by inhibition, so as to not assume the control of action/output at retrieval (Marsh et al., 2013) and impair retrieval of category exemplars with low output dominance.

That the semantic properties of irrelevant speech do not interfere with serial order is perhaps due in part to the use of digits and letters that are by design impoverished in terms of their semantic content. For the few studies in which the semantic properties of speech have produced small amounts of disruption, the target items are typically also words (Marsh et al., 2008; Neely & LeCompte, 1999). Here it is likely that long-term memory factors such as lexicality and semantic similarity that are known to contribute to immediate serial recall performance through co-opting of the speech production process (Hulme, Maughan, & Brown, 1991; Saint-Aubin & Poirier, 1999) could indeed become susceptible to interference due to the semantic processing of irrelevant speech. Moreover, we suspect that the only study to show that meaningful speech produces more distraction than meaningless speech (LeCompte, Neely, & Wilson, 1997) wherein digits were used as targets, may be more plausibly attributable to the attentionally capturing nature of the irrelevant words ("hey, you, me, no") or a lack of acoustic complexity within the control condition: A time-reversed version of the four spoken words. In the context of these few tokens (as compared to longer passages of reversed speech), reversing speech may produce less abrupt onsets of the transient parts of acoustic signal (stop bursts and vowel onsets) and other spectro-temporal changes, in comparison with its forward speech counterpart, which could have reduced the changing-state (i.e. acoustic) properties of the acoustic signal, thereby reducing the disruption of serial recall. In short, the reversed speech signal may contain less changing-state, and thus interfere less with serial recall, in comparison with the regular speech.

If it is processing of the lexical-item that produces disruption, rather than inter-lexical processing, this may be seen by some as rather blurring the distinction between accounts that assume that interference arises at the level of the item rather than at the level of the sequence, for example. Although processing lexical-item identity appears important here, it is the consequence that this processing has on more high-order processing (assembly of categorically organized retrieval plans for output) that appears to be a strong determinant of distraction. Our position here is shaped by the finding that unrelated speech produces just as much disruption as related speech to the number of categories recalled. Therefore, that the distractors have lexical-item identity would appear to be the sole cause of the impairment in organization of target responses, at least as indexed by loss of categories in the irrelevant speech conditions. However, between-sequence semantic similarity has an additional effect on organization, as is reflected in the disruption to the organization of output according to those semantic categories recalled (clustering). This outcome is consistent with the finding that the lexical properties of unrelated speech also impairs recall of lists wherein all members are derived from a single semantic category, with additional disruption arising from speech that is related to the to-be-recalled list (Marsh et al., 2008). In this setting, the non-lexical properties of irrelevant sound do not produce disruption. Words, but not non-words nor sine-wave speech, produces disruption unlike in the case of serial recall (Tremblay, Nicholls, Alford & Jones, 2000). Even in the context of recalling lists comprising single semantic categories, the semantic richness of the items as well as the longer list-length, promotes semantic-based organization processes. However, such organizational processes cannot be directly observed with recall of single-category lists since traditional

measures of mnemonic organization such as category recall and clustering measures cannot be computed (Marsh et al., 2009).

Conclusion

The semantic auditory distraction phenomenon can be explained without assuming a bespoke memory store. Rather, it seems like involuntary analysis of the meaning of unattended speech, at the lexical level, disrupts on-going semantic processes that are involved in the organization and preparation of behavioural output. Disruption takes place because dominant (or 'overlearned') responses receive activation and compete for the control of action.

References

Allen, M. (1968). Rehearsal strategies and response cueing as determinants of

    organization in free recall. *Journal of Verbal Learning & Verbal Behavior*,

    7, 58-63.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the

    mechanisms of forgetting. *Journal of Memory & Language*, *49*, 415–445.

Baddeley, A. D., Eldridge, M., Lewis, V., & Thomson, N. (1984). Attention and

    retrieval from long-term memory. *Journal of Experimental Psychology:*

    General, *113*, 518-540.

Beaman, C. P. (2004). The irrelevant sound effect revisited: What role for working

    memory capacity? *Journal of Experimental Psychology: Learning, Memory,*

    & Cognition, *30*, 1106-1118.

Bell, R., Buchner, A., & Mund, I. (2008). Age-related differences in irrelevant-speech

    effects. *Psychology & Aging, 23*, 377-391.

Bell, R., Mund, I., & Buchner, A. (2011). Disruption of short-term memory by distractor

    speech: Does content matter? *Quarterly Journal of Experimental Psychology, 64,*

    146-168.

Bell, R., Röer, J. P., Dentale, S., & Buchner, A. (2012). Habituation of the irrelevant

    sound effect: Evidence for an attentional theory of short-term memory disruption.

    Journal of Experimental Psychology: Learning, Memory, & Cognition, 38, 1542-

    1557.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly

    arranged associates. *Journal of General Psychology, 49*, 229-240.

Bousfield, W. A., & Puff, C. R. (1964). Clustering as a function of response dominance. *Journal of Experimental Psychology, 67*, 76-79.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning & Verbal Behavior*, *8*, 323-343.

Broadbent, D. (1983). Recent advances in understanding performance in noise. In G. Rossi (Ed.). *Noise as a public health problem*: *Proceedings of the Fourth International Congress* (pp. 719-738). Milano: Edizioni Techniche a Cura del Centro Ricerche e Studi Amplifon.

Buchner, A., Irmen, L., & Erdfelder, E. (1996). On the irrelevance of semantic information in the 'irrelevant speech' effect. *Quarterly Journal of Experimental Psychology, 49A*, 765-779.

Buchner, A., Mehl, B., Rothermund, K., & Wentura, D. (2006). Artificially induced valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory & Cognition, 34,* 1055-1062.

Burns, D. J., & Brown, C. A. (2000). The category access measure of relational processing. *Journal of Experimental Psychology*: *Learning*, *Memory*, *& Cognition*, *26*, 1057-1062.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. (1966). Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology, 71*, 858-866.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.

Evans, G. W., & Johnson, D. (2000). Stress and open-office noise. *Journal of Applied Psychology, 85*, 779-783.

Frankel, F., & Cole, M. (1971). Measures of category clustering in free recall. *Psychological Bulletin, 76*, 39-44.

Gathercole, S. E., & Baddeley, A. D. (1993). Working memory and language. Hillsdale, NJ: Erlbaum.

Gronlund, S. D., & Shiffrin, R. M. (1986). Retrieval strategies in recall of natural categories and categorized lists. *Journal of Experimental Psychology*: *Learning*, *Memory*, *& Cognition*, *12*, 550-561.

Hudson, R. L. (1969). Category clustering for immediate and delayed recall as a function of recall cue information and response dominance variability. *Journal of Experimental Psychology, 82*, 575-577.

Hughes, R. W., & Jones, D. M. (2003). A negative order-repetition priming effect: Inhibition of order in unattended sequences? *Journal of Experimental Psychology*: *Human Perception & Performance*, *29*, 199-218.

Hughes, R. W., & Jones, D. M. (2005). The impact of order incongruence between a task-irrelevant auditory sequence and a task-relevant visual sequence. *Journal of Experimental Psychology*: *Human Perception & Performance*, *31*, 316-327.

Hughes, R. W., Vachon, F., & Jones, D. M. (2005). Auditory attentional capture during serial recall: Violations at encoding of an algorithm-based neural model? *Journal of Experimental Psychology*: *Learning, Memory, & Cognition*, *31*, 736-749.

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by

changing and deviant sounds: Support for a duplex mechanism account of
auditory distraction. *Journal of Experimental Psychology: Learning, Memory, &
Cognition, 33*, 1050-1061.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and
distinctiveness. *Journal of Memory & Language, 32*, 421-445.

Hygge, S., Boman, E., & Enmarker, I. (2003). The effects of road traffic noise and
meaningful irrelevant speech on different memory systems. *Scandinavian Journal
of Psychology, 44,* 13-21.

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 & Cognition, 19,
369-381.

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech
to short-term memory: The role of changing-state. *Quarterly Journal of
Experimental Psychology*, *44A*, 645-669.

Jones, D. M., Marsh, J. E., & Hughes, R. W. (2012). Retrieval from memory: Vulnerable
or inviolable? *Journal of Experimental Psychology: Learning, Memory, &
Cognition, 38*, 905-922.

Jones, D. M., Miles, C., & Page, J. (1990). Disruption of proof-reading by irrelevant
speech: Effects of attention, arousal or memory? *Applied Cognitive
Psychology, 4*, 89-108.

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

Macken, W. J., Tremblay, S., Houghton, R. J., Nicholls, A. P., & Jones, D. M. (2003).

    Does auditory streaming require attention? Evidence from attentional

    selectivity in short-term memory. *Journal of Experimental Psychology*:

    Human Perception & Performance, *29*, 43-51.

Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in

    music and language. *Journal of Cognitive Neuroscience, 23*, 2232-2247.

Marsh, J. E., Beaman, C. P., Hughes, R. W., & Jones, D. M. (2012). Inhibitory control in

    memory: Evidence for negative priming in free recall. *Journal of Experimental*

    Psychology: Learning, Memory, & Cognition, 38, 1377-1388.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008). Auditory distraction in semantic

    memory: A process-based approach. *Journal of Memory & Language, 58*, 682-

    700.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content,

    determines semantic auditory distraction. *Cognition, 110*, 23-38.

Marsh, J. E., & Jones, D. M. (2010). Cross modal distraction by background speech:

    What role for meaning? *Noise & Health*, *12*, 210-216.

Marsh, J. E., Sörqvist, P., Beaman, C. P., & Jones, D. M. (2013). Auditory distraction

    eliminates retrieval induced forgetting: implications for the processing of

    unattended sound. *Experimental Psychology*, *60*, 368-375.

Marsh, J. E., Sörqvist, P., Halin, N., Nöstl, A., & Jones, D. M. (2013). Auditory

    distraction compromises random generation: Falling back into old habits?

    Experimental Psychology, 60, 1-14.

Marsh, J. E., Vachon, F., & Jones, D. M. (2008). When does between-sequence phonological similarity produce irrelevant sound disruption? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34*, 243-248.

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory & Language, 27,* 382-398.

Mood, A. M. (1940). The distribution theories of runs. *Annals of Mathematical Statistics,11*, 367-392.

Moscovitch, M. (1994). Cognitive resources and dual-task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. *Neuropsychology, 8*, 524-534.

Mulligan, M. W. (1999). The effects of perceptual interference at encoding on organization and order: Investigating the roles of item-specific and relational information. *Journal of Experimental Psychology*: *Learning, Memory, & Cognition, 25*, 54-69.

Murphy, M. D. (1979). Measurement of category clustering in free recall. In C. R. Puff (Ed.). *Memory organization and structure* (pp. 51-83). San Diego, CA: Academic Press.

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

Neely, C. B., & LeCompte, D. C. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory & Cognition, 27*, 37-44.

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence

for facilitatory and inhibitory processes. *Memory & Cognition*, *4*, 648-654.

Neely, J. H., & Kahan, T. A. (2001). Is semantic activation automatic: A critical re-evaluation. In H. L. Roediger, J. S. Nairne, I. Neath & A. M. Surprenant (Eds.), *The nature of remembering*: *essays in honor of Robert G. Crowder*. Washington DC: American Psychological Association.

Neumann, O. (1996). Theories of attention. In O. Neumann and A.F. Sanders (Eds.), *Handbook of perception and action*, (Vol. 3, pp. 389-446). London: Academic Press.

Oswald, C. J. P., Tremblay, S., & Jones, D. M. (2000). Disruption of comprehension by the meaning of irrelevant sound. *Memory*, *8*, 345-350.

Parmentier, F. B. R. (2008). Towards a cognitive model of distraction by auditory novelty: The role of involuntary attention capture and semantic processing. *Cognition, 109,* 345-362.

Sheffert, S. M., Pisoni, D. B., Fellowes, J. M., & Remez, R. E. (2002). Learning to recognize talkers from natural, sinewave, and reversed speech samples. *Journal of Experimental Psychology*: *Human Perception and Performance*, *28*, 1447-1469.

Shuell, T. J. (1968). Retroactive inhibition in free-recall learning of categorized lists. *Journal of Verbal Learning and Verbal Behavior*, *7*, 797–805.

Shuell, T. J. (1969). Clustering and organization in free recall. *Psychological Bulletin, 72*, 353-374.

Sirotin, Y. B., Kimball, D. R., & Kahana, M. J. (2005). Going beyond a single list: Modeling the effects of prior experience on episodic free recall. *Psychonomic Bulletin & Review*, *12*, 787-805.

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology*, *81*, 557-560.

Smith, A. P., Jones, D. M., & Broadbent, D. E. (1981). The effects of noise on recall of categorized lists. *British Journal of Psychology, 72*, 299-316.

Sörqvist, P. (2010). Effects of aircraft noise and speech on prose memory: What role for working memory capacity? *Journal of Environmental Psychology, 30,* 112-118.

Sörqvist, P., Marsh, J. E., & Jahncke, H. (2010). Hemispheric asymmetries in auditory distraction. *Brain & Cognition, 74,* 79-87.

Sörqvist, P., Nöstl, A., & Halin, N. (2012). Disruption of writing processes by the semanticity of background speech. *Scandinavian Journal of Psychology, 53,* 97-102.

Sundstrom, E., Town, J. P., Rice, R. W., Osborn, D. P., & Brill, M. (1994). Office noise, satisfaction, and performance. *Environment & Behavior, 26*, 195-222.

Tulving, E. (1968). Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.). *Verbal behavior and general behavior theory* (pp. 2-36). Englewood Cliffs, N. J.: Prentice Hall.

Tun, P. A., O' Kane, G., & Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychology & Aging*, *17*, 453-467.

Underwood, G. (1977). Contextual facilitation from attended and unattended messages. *Journal of Verbal Learning & Verbal Behavior, 16,* 99–106.

Underwood, G. (1981). Lexical recognition of embedded unattended words: Some implications for reading processes. *Acta Psychologica, 47*, 267-283.

Underwood, G., & Everatt, J. (1996). Automatic and controlled information processing:

The role of attention in the processing of novelty. In O. Neumann and A.F.

Sanders (Eds.), Handbook of perception and action, (Vol. 3, pp. 185-221).

London: Academic Press.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An

updated and expanded version of the Battig and Montague (1969) norms. *Journal*

of Memory & Language, *50*, 289-335.

Wallis, W. A., & Roberts, H. V. (1957). *Statistics: A new approach*. New York: Free

Press of Glencoe.

Weist, R. M. (1972). The role of rehearsal: Recopy or reconstruct? *Journal of Verbal*

Learning & Verbal Behavior, *11*, 440-450.

Wetherick, N. E. (1975). The Role of Semantic Information in Short-Term Memory.

Journal of Verbal Learning and Verbal Behavior, *14*, 471-480.

Witterseh, T., Wyon, D. P., & Clausen, G. (2004). The effects of moderate heat stress and

open-plan office noise distraction on SBS symptoms and on the performance of

office work. *Indoor Air, 14*, 30-40.

Yoon, C., Feinberg, F., Hu, P., Gutchess A. H., Hedden, T., Chen, H., Jing, Q., Cui,

Y., & Park, D. C. (2004). Category norms as a function of culture and age:

Comparisons of item responses to 105 categories by American and Chinese

adults. *Psychology & Aging*, *19*, 379-393.

Zhang, Y., Li, P., Piao, Q., Liu, Y., Huang, Y., & Shu, H. (2013). Syntax does not

necessarily precede semantics in sentence processing: ERP evidence from

Chinese. *Brain & Language, 126*, 8-19.

Author note

Table 1

Mean recall and clustering measure as a function of irrelevant sound condition in

Experiment 1

Sound condition	Study only		Test only	
	M	*SE*	*M*	*SE*
Category-exemplars correctly recalled				
Quiet	15.89	.52	15.35	.55
Reversed speech	15.20	.55	14.67	.54
Forward speech	13.99	.54	13.58	.56
Exemplars per category correctly recalled				
Quiet	4.22	.11	4.08	.12
Reversed Speech	4.0	.12	3.92	.11
Forward Speech	3.79	.12	3.74	.12
Categories correctly recalled				
Quiet	3.78	.05	3.74	.04
Reversed Speech	3.78	.04	3.71	.05
Forward Speech	3.67	.05	3.59	.05
Z scores				
Quiet	3.57	.21	3.46	.22
Reversed Speech	3.32	.23	3.49	.20
Forward Speech	3.04	.22	3.12	.21

The table above has a header "Locus of sound presentation" spanning the Study only and Test only columns.

Table 2

Mean recall and clustering measure as a function of irrelevant sound condition in

Experiment 2

	Sound condition							
	Quiet		Text-speech		3rd-order approximation		5th-order approximation	
Dependent measure	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Category-exemplars correctly recalled	17.24	.76	14.95	.74	14.15	.58	14.63	.66
Exemplars per category correctly recalled	4.40	.17	4.06	.16	3.89	.17	3.95	.15
Categories correctly recalled	3.90	.03	3.66	.06	3.66	.06	3.69	.06
Z scores	4.38	.21	3.61	.25	3.35	.24	3.52	.22

Table 3

Mean recall and clustering measures as a function of irrelevant sound condition and dominance condition (HH: high dominance to-be-recalled items and high dominance to-be-ignored items; HL: high dominance to-be-recalled items and low dominance to-be-ignored items; LH: low dominance to-be-recalled items and high dominance to-be-ignored items; and LL: low dominance to-be-recalled items and low dominance to-be-ignored items) in Experiment 3

| | Dominance condition | | | | | | | |
| | HH | | HL | | LH | | LL | |
Sound condition	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Category-exemplars correctly recalled								
Unrelated speech	17.68	.58	15.69	.50	12.18	.72	12.28	.72
Related speech	15.70	.61	14.86	.41	8.83	.41	12.25	.77
Exemplars per category correctly recalled								
Unrelated speech	4.44	.15	3.83	.18	3.31	.15	3.11	.18
Related speech	3.94	.15	3.61	.16	2.42	.10	3.14	.20
Categories correctly recalled								
Unrelated speech	3.99	.01	4.00	.00	3.66	.09	3.95	.02
Related speech	3.98	.02	3.98	.01	3.64	.05	3.91	.03
Z scores								
Unrelated speech	4.59	.29	3.00	.27	2.69	.29	2.27	.31
Related speech	4.15	.28	3.12	.20	1.44	.16	2.17	.28