THE IRRELEVANT SPEECH EFFECT AND THE LEVEL OF INTERFERENCE IN AGING

Pascal W. M. Van Gerven,
Department of Neurocognition, Faculty of Psychology, Maastricht University, Maastricht, The Netherlands

Willemien A. Meijer
Department of Psychiatry and Neuropsychology, Faculty of Medicine, Maastricht University, Maastricht, The Netherlands

Annemiek Vermeeren
Department of Neurocognition, Faculty of Psychology, Maastricht University, Maastricht, The Netherlands

Eric F. Vuurman
Department of Psychiatry and Neuropsychology, Faculty of Medicine, Maastricht University, Maastricht, The Netherlands

Received 25 August 2005; accepted 18 December 2005.

The authors thank Mieke Janssen and Céline Penando for testing the participants. Anita van Oers is gratefully acknowledged for adapting the E-Prime tasks and Johan Gielissen and Iemke Horemans are acknowledged for their technical advice on recording the background speech. Finally, the authors thank Wouter Janssen for recording the Dutch, and Sergey Ripinsky for recording the Russian background stimuli.

Address correspondence to Pascal W. M. Van Gerven, Department of Neurocognition, Faculty of Psychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. E-mail: p.vangerven@psychology.unimaas.nl
In general, no disproportionate detrimental effects of irrelevant background speech on cognition are found in aging individuals, although this is predicted by the inhibitory view of aging. This may be due to the nature of the primary task (most studies involve a verbal learning task) or the cognitive level at which irrelevant speech interferes with this task. In this study, the irrelevant speech effect on a numeric working memory task was investigated among 20 young (M = 21.8 years) and 20 older (M = 68.1 years) native Dutch individuals. Level of interference (LOI) was manipulated by presenting white noise (no interference), Russian words (low interference), Dutch words (phonological interference), and Dutch numbers (semantic interference) in the background. Results showed that reaction time increases as a function of LOI relative to silence, whereas accuracy remains unaffected. However, no interaction between LOI and age group was found, which suggests that the elderly were not disproportionally affected by an increased level of interference. These results are discussed in the light of the inhibitory view of aging.

Verbal noise is omnipresent in modern, densely populated societies. Its influence is potentially harmful to cognitive performance in professional environments, such as open-plan offices (e.g., Banbury & Berry, 1998). In the laboratory, effects of verbal noise are usually studied by presenting participants, who are engaged in some visually presented cognitive task, with irrelevant background babble, narration, conversation, or random words. Starting from Baddeley's (1992, 2003) conception of working memory, the idea is that the visually presented task material is transformed into a phonological representation (i.e., internal speech), which meets up with the background speech in the phonological loop. It is generally found that the background speech interferes with cognitive performance in that it leads to higher response times and decreased accuracy. This effect is known as the unattended or irrelevant speech effect (e.g., Baddeley & Salamé, 1986; Miles, Jones, & Madden, 1991; Salamé & Baddeley, 1982).

The disruptive effect of speech seems distinct from the effect of nonverbal noise, such as white noise, which is generally found not to impede cognitive performance. However, a range of studies aimed at mimicking certain characteristics of speech in nonverbal noise, such as changes in pitch, have yielded comparable disruptive effects (e.g., Jones, 1995; Jones & Macken, 1993; Tremblay, Nicholls,
Alford, & Jones, 2000). The presumed mechanism underlying this phenomenon has come to be known as the “changing-state hypothesis,” which negates the unique role of speech in accounting for the adverse effects of background noise on cognition (Bridges & Jones, 1996; Jones, Alford, Bridges, Tremblay, & Macken, 1999; Jones & Macken, 1995; Jones, Madden, & Miles, 1992). The changing-state hypothesis is further supported by studies yielding comparable effects of verbal and nonverbal noise if the latter contains some acoustic variation (e.g., Hygge, Boman, & Enmarker, 2003). Results of other studies, however, do not support the changing-state hypothesis, because they yield larger effects for verbal than for nonverbal noise (e.g., LeCompte, Neely, & Wilson, 1997; Neely & LeCompte, 1999; Salamé & Baddeley, 1989). Whether speech is unique in inducing interference effects or not, the bottom line is that certain acoustic characteristics of verbal noise—irrespective of whether they can be mimicked or not—certainly play a key role in the irrelevant speech effect.

Besides acoustic or phonological interference, irrelevant speech is likely to cause semantic interference. Research has shown that the strength of semantic interference largely depends on the meaningfulness of the background speech. For example, Oswald, Tremblay, and Jones (2000) found cognitive performance to be more susceptible to meaningful than to meaningless background speech. Furthermore, the magnitude of semantic interference effects seems to be determined by the resemblance between the irrelevant information that is conveyed by the speech and the information that is processed during the task at hand. A study by Neely and LeCompte (1999), for example, revealed that irrelevant speech conveying words that were strongly associated with to-be-learned verbal material impeded learning of that material more than irrelevant speech conveying unrelated words. Other studies have suggested, however, that semantics does not play a crucial role in the irrelevant speech effect (e.g., Buchner, Irmen, & Erdfelder, 1996).

Relatively little research on the irrelevant speech effect has been conducted with older participants. An early study by Molander and Bäckman (1990) with very few participants (six young and six older adults) revealed that motor performance of older people is more affected by verbal than nonverbal noise, whereas motor performance of younger adults is equally affected by the two types of noise. Meijer, De Groot, Van Boxtel, Van Gerven, and Jolles (2006) found episodic memory of older adults to be more affected by irrelevant speech only after decreasing the interstimulus interval in a verbal learning task. In most studies, however, no interaction was found between noise and age. For example, Enmarker (2004) found irrelevant speech to impede memory performance (both episodic and semantic) and
attention, but this effect did not differ between a group of young (aged 35 to 45 years) and a group of older adults (aged 55 to 65 years). Comparable effects on episodic memory were found by Rouleau and Belleville (1996). Also if relevant and irrelevant information are both presented orally, older individuals do not seem to me more affected than their younger counterparts. Li, Daneman, Qi, and Schneider (2004), for example, found that younger and older adults are equally disturbed by natural background conversation if they are instructed to repeat orally presented sentences. Boman, Enmarker, and Hygge (2005) extended the absence of an age-specific irrelevant speech effect to a particularly broad age range, which included individuals between 13 and 65 years of age. Even in older adults with Alzheimer’s disease, episodic memory performance does not seem to be more impaired by background speech than in neurologically healthy older and younger controls (Belleville, Rouleau, Van der Linden, & Collette, 2003). Finally, the apparent nonexistence of disproportionate noise effects in the elderly cannot be explained by age-related hearing disorders (Beaman, 2005; Murphy, McDowd, & Wilcox, 1999).

The above findings are peculiar in the light of the inhibitory view of cognitive aging (Hasher, Stoltzfus, Zacks, & Rypma, 1991; Hasher & Zacks, 1988; McDowd, 1995; Stoltzfus, Hasher, Zacks, Ulivi, & Goldstein, 1993), which contends that older people are more affected by the negative influence of distracting stimuli than young people, because of an age-related decline of inhibitory control. The aforementioned pattern of results becomes even more peculiar if we consider the age-related decline of the frontal lobes (e.g., Tisserand & Jolles, 2003), which are associated with both inhibitory control (e.g., Braver & Barch, 2002) and the resistance to irrelevant speech in working memory (e.g., Gisselgård, Petersson, & Ingvar, 2004).

A first possible reason for the absence of a noise by age interaction effect is that all studies were aimed at learning, recall, and recognition of digits or words. No studies could be retrieved addressing the effect of irrelevant speech on central executive functions of working memory. Although verbal learning is known to gradually decline with aging (Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005), reduced working memory performance—especially under conditions that require high levels of cognitive control—is generally regarded as one of the most robust manifestations of age-related cognitive decline (Braver & Barch, 2002) and is therefore likely to be particularly vulnerable to irrelevant background speech. The current study will focus on the Star Counting Test (SCT; De Jong & Das-Smaal, 1990, 1995), which was originally designed to probe the central executive of working memory in children. The SCT requires the participant
to count stars on a display. At certain points, the counting direction must be changed from forward to backward or vice versa. Switching of counting direction is assumed to be controlled by the central executive. Storage, maintenance, and rehearsal of the count are assumed to take place in the phonological loop. Irrelevant speech is likely to interfere with the current count in the phonological loop (see Logie & Baddeley, 1987), dependent on the effectiveness of inhibitory control by the central executive.

A second possible reason for the absence of a disproportionate aging effect in the aforementioned studies is that the cognitive level of interference (e.g., perceptual, phonological, semantic) was not high enough to elicit a disproportionate irrelevant speech effect in the elderly. The study by Enmarker (2004), for example, entailed only two levels of interference (with silence as the control condition): road traffic noise and familiar irrelevant speech (Swedish). The studies by Belleville and colleagues (Belleville et al., 2003; Rouleau & Belleville, 1996) involved three levels of interference: white noise, unfamiliar irrelevant speech (Romanian), and familiar irrelevant speech (French). In the current study, an extra level of interference was added by increasing the similarity between the contents of the irrelevant speech and the information to be processed in the SCT (cf. Neely & LeCompte, 1999). This was done by confronting the participants with verbally presented random numbers. The similarity between irrelevant numbers and the running count during the SCT is likely to cause interference at a lexicosemantic level (Logie & Baddeley, 1987), which would be reflected by reduced speed and accuracy.

In sum, the current study involved five conditions: (1) silence, (2) white noise, (3) Russian words, (4) Dutch words, and (5) Dutch numbers. These conditions were supposed to gradually increase the level of interference. Silence served as a control condition. White noise was not expected to cause any interference. Russian words were expected to cause some interference at the phonological level, because they are recognized as speech and thus convey speech-specific “changing state” characteristics (apart from that, Russian is unrelated to Dutch, because it stems from the Slavic language family, whereas Dutch stems from the Germanic language family). Dutch words were expected to at least cause interference at the phonological level. Dutch numbers, finally, were expected to cause interference up to the semantic level. Furthermore, based upon the inhibitory view of cognitive aging, it was hypothesized that the level of interference interacts with age group (young, old). More specifically, performance of the older participants was expected to be disproportionately impaired by the higher levels of interference, especially in the Dutch.
numbers condition, which involved an increased semantic similarity between relevant and irrelevant information.

**METHODS**

**Participants**

Twenty young \((M = 21.8\text{ years}, SD = 2.0; 10\text{ men and 10 women})\) and 20 older adults \((M = 68.1\text{ years}, SD = 4.5; 10\text{ men and 10 women})\) volunteered to participate in the study. The young participants were psychology students at Maastricht University. The older participants were recruited from a participant pool of the Maastricht Aging Study (MAAS; Jolles, Houx, Van Boxtel, & Ponds, 1995). They were selected to participate if they were in good physical and psychological health, had at least received college-level education, were native Dutch speakers, and had no knowledge of the Russian language. All participants had normal or corrected-to-normal vision and no hearing disabilities. They received €10 or course credits for their participation.

To get an impression of the participants’ cognitive speed and susceptibility to interference, the Stroop Color-Word Test (SCWT; Houx, Jolles, & Vreeling, 1993) was administered. On this test, the elderly participants were significantly slower than the young participants at reading color names (Card 1) and naming color patches (Card 2), \(t(38) = -2.70, p = .01\). The older participants were also more vulnerable to Stroop interference (i.e., time on Card 3 [incongruent color naming] minus the average time on Cards 1 and 2) than the young, \(t(38) = -3.40, p < .01\). This pattern of results is in line with cognitive slowing and inhibitory decline theories of aging (e.g., Hasher & Zacks, 1988; Salthouse, 1996).

**Materials**

**Star Counting Test (SCT)**

The Star Counting Test was originally devised by De Jong and Das-Smaal (1995). The current version of the task was programmed in E-Prime version 1.1 (Psychology Software Tools, Pittsburgh, PA). In a typical trial, a display consisted of nine lines of up to six unevenly spaced asterisks (see Figure 1). The task of the participant was to count the stars from left to right and row by row (i.e., from top to bottom), starting from a random number that was indicated in the
upper left corner of the display. Initially, the participant was instructed to count forward. If a minus sign was encountered in the star field, however, subsequent stars had to be counted backward. If a plus sign was encountered, stars had to be counted forward again. After arriving at the last star, the counting result had to be compared to a probe number in the lower right corner of the display. If this number was equal to the counting result, the participant had to press a green key labeled “yes”; if the number was unequal to the counting result, which is the case in Figure 1 (probe is 37, correct counting result is 38), a red key labeled “no” had to be pressed. In the negative trials, the probe number deviated only one from the correct counting result.

**Noise and Irrelevant Speech Sources**

Noise and irrelevant speech was presented through a headphone, which was connected to a CD player. The output volume of the CD player was set to a fixed level of approximately 70 dB(A). A first track contained white noise, which represented the nonverbal noise condition. A second track contained random Russian words, which were pronounced by a native Russian male. The Russian words were
translations of a third track of random Dutch words. This was a
selection of words with a frequency of occurrence ranging between
40 and 100 per million words, which was drawn from the CELEX
lexical database (Max Planck Institute for Psycholinguistics, Nijme-
gen, The Netherlands). These words were pronounced by a native
Dutch male. A fourth and final track contained random Dutch num-
bers between 1 and 110. This range was similar to the counting range
of the SCT. The numbers were pronounced by the same Dutch
male who pronounced the Dutch random words. All irrelevant-speech
stimuli were articulated at a rate of about one item per 1.5 s.

**Design and Procedure**

Age group (young, old) was the independent between-group variable.
Level of interference (LOI: silence, white noise, Russian words,
Dutch words, Dutch numbers) was the independent within-groups
variable. This yielded a $2 \times 5$ full-factorial design. Dependent
variables were response time (in seconds) and accuracy (number of
correct responses).

The experiment was performed in individual sessions. The partici-
pant received on-screen instructions, but was allowed to ask ques-
tions to the experimenter. The experiment was preceded by two
practice trials: one trial had to be evaluated as correct (i.e., star
counting result was equal to probe number) and one as incorrect
(i.e., star counting result was unequal to probe number). During
the practice trials, the participant was requested to count aloud to
make sure that the instructions were properly understood.

The order of the LOI conditions was counterbalanced according to
a $5 \times 5$ Latin square. To prevent practice effects, five different ver-
sions of the SCT were used, each of which consisted of nine trials
(i.e., nine star fields). The order of these test versions was counterba-
lanced as well. However, to prevent a confound between test version
and LOI, a “reversed” $5 \times 5$ Latin square was applied to the test ver-
sions. That is, where LOI conditions were shifted from left to right in
each consecutive row, test versions were shifted from right to left. In
this way, every condition was combined with every test version.

**RESULTS**

All data were analyzed with a $2 \times 5$ (age group $\times$ LOI) repeated-
measure analysis of variance (ANOVA). In the case of a significant
effect of LOI, repeated contrasts (i.e., comparison of adjacent LOI
conditions) and simple contrasts (i.e., comparison of the silence con-
condition with all other LOI conditions) were analyzed, which yielded a total of seven unique comparisons. Alpha levels were corrected according to the Bonferroni-Holm procedure, which is conservative, but appropriate if relatively many comparisons are involved. Following this procedure, the $p$ values corresponding to the contrasts were first ordered from the smallest to the largest. Subsequently, the smallest $p$ value was compared to an alpha level of .05 divided by the number of comparisons, $k$. The next smallest $p$ value was compared to alpha divided by $k-1$, the next by $k-2$, and so on. The procedure was stopped as soon as a contrast was non-significant, implying that subsequent contrasts were non-significant as well.

**Accuracy**

Accuracy was determined as the number of correct responses (range: 0 to 9). Mean accuracy scores are given in Table 1. The average accuracy was high for both age groups (i.e., at least 6 out of 9 correct trials). The older participants were generally less accurate than the young participants, $F(1, 38) = 9.80$, $MSE = 0.56$, $p < .01$. Accuracy was not significantly affected by LOI in either of the age groups, $F(4, 152) = 1.31$, $MSE = 1.23$, $p = .27$. Furthermore, age group did not interact with LOI, $F(4, 152) = 1.37$, $p = .25$. No further contrast analyses were performed on these data.

**Response Time**

Only response times corresponding to correct responses were included in the analysis. A graphical representation of mean response times (in seconds) can be found in Figure 2. The older participants appeared to be slightly slower than the young participants, but this
difference only approached significance, $F(1, 38) = 3.35, MSE = 45.13, p = .075$. Most importantly, there was a significant main effect of LOI, $F(4, 152) = 3.39, MSE = 12.64, p = .01$, indicating that response latencies increased with the level of interference (see Figure 2). The effect of LOI did not interact with age group ($F < 1$), which suggests that the elderly participants were not more affected by an increased level of interference than their young counterparts. Possibly, the absence of this interaction was partly due to a lack of statistical power. Therefore, a post-hoc power analysis was performed using the GPOWER software package (Faul & Erdfelder, 1992). Starting from a medium effect size of .15 and an alpha level of .05, this yielded a power of .97. An “a priori” power analysis revealed that a power of around .90 is already attainable with 15 participants per age group ($N = 30$).

Following the Bonferroni-Holm procedure, the contrast between the silence and the Dutch numbers condition revealed the highest level of significance and thus was the first contrast to be evaluated. Starting at an alpha level of .007, there was a significant main effect of LOI, $F(1, 38) = 11.91, MSE = 27.45, p = .001$, indicating that participants were significantly slower in the Dutch numbers condition than in the silence condition. This effect did not interact with age group ($F < 1$). The second contrast to be evaluated using an alpha level of .008 was that between the Dutch words and the Dutch numbers condition. This analysis did neither reveal a significant main effect of LOI, $F(1, 38) = 3.46, MSE = 28.23, p = .07$, nor an

![Figure 2. Mean response time (in seconds) as a function of the level of interference (LOI). Error bars indicate one standard error of the mean.](image-url)
interaction with age group \((F < 1)\). No further contrast analyses were performed.

**DISCUSSION**

The main hypothesis of the current study was that, relative to young adults, counting performance of older adults is disproportionately impaired by irrelevant speech if there is interference at relatively high levels of processing. A study by Rouleau and Belleville (1996) did not reveal such a disproportionate effect using irrelevant speech that was relatively unrelated to the digits in a serial recall task. These authors suggested, however, that the elderly might show impaired performance if irrelevant speech interferes at a lexicosemantic level. In the current study, semantic interference was induced by presenting the participants with irrelevant numbers that were in the same range as the numbers that were processed in the Star Counting Task (SCT). However, even this ultimate level of interference did not result in a disproportionate performance impairment in the elderly. Only a main effect of the level of interference \((LOI)\) was found, which was largely attributable to the irrelevant numbers condition relative to the silence condition. The main effect was only obtained for reaction time, however, not for accuracy. This might indicate that performing the SCT under irrelevant speech conditions does not so much involve an increased risk of losing count, but rather an increased effort to maintain the count in the phonological loop. For example, it is conceivable that the frequency of rehearsal in the phonological loop increases under verbal noise, which inevitably leads to longer response times.

The results of the present study do not support Rouleau and Belleville’s (1996) idea that lexicosemantic inhibition is disrupted by aging, while phonological inhibition is relatively intact. Instead, the current data suggest that both phonological and semantic inhibition are relatively intact in older adults. There may be two possible reasons why this result was obtained. First, the older participants were highly educated, which might be a protective factor against cognitive aging (see, e.g., Stern, 2002). Second, the counting task might address an “overlearned” skill, which is relatively insensitive to interference by background speech. The latter explanation is disputable, however, because the SCT did not only involve forward counting, but also backward counting, which might be less well trained. Moreover, switches of counting direction made the SCT particularly complex. Nevertheless, an interaction between age and the level of interference might emerge if task complexity is systematically increased. Meijer et al. (2006), for instance, found greater disruptive effects of
background speech on verbal learning in older than in young individuals if they decreased the time interval between words in the encoding phase of a verbal learning task.

A possible age-independent confound in the current study might be the "stimulus density" in the irrelevant speech conditions. Buchner, Steffens, Irmen, and Wender (1998) found the detrimental effects of irrelevant verbal background speech on counting performance to be dependent on the number of syllables it contained per time unit. That is, the higher the syllable frequency, the stronger the irrelevant speech effect. Consequently, Buchner et al. found no differential effects of several irrelevant speech conditions (i.e., words, random numbers, distant numbers, and adjacent numbers) after controlling for the number of syllables per time unit. In the present study, the number of syllables per time unit was not controlled for. Thus, the Dutch numbers condition might have contained a relatively high number of syllables regarding the range from which they were drawn (1 to 110).

Although the irrelevant speech effect in the Dutch numbers condition relative to the silence condition might have been caused by a relatively high syllable density, the effect did not differ between the age groups. In that respect, the current study is consistent with most previous studies on irrelevant speech and aging (e.g., Beaman, 2005; Enmarker, 2004; Rouleau & Belleville, 1996). Still, these results are remarkable in the light of the inhibitory view of aging, which would obviously predict a disproportionate irrelevant speech effect in the elderly. Two explanations can be put forward that potentially shed more light on this issue.

First, there is the role of cognitive effort. Although the elderly's performance was not disproportionally impaired by irrelevant speech, they could have invested more mental effort to maintain their performance level than their young counterparts. Considering common cognitive aging phenomena, such as a changing speed-accuracy tradeoff, there is ample reason to assume that older people need to invest more effort to achieve a comparable level of performance as young people. The greater susceptibility of the older participants to Stroop interference in the current study might indeed indicate that they had to devote more effort to successfully ignore the irrelevant speech. Increased cognitive effort in the elderly might become visible as increased brain activity by means of functional MRI or other imaging techniques. More specifically, the elderly might show increased activity in the dorsolateral prefrontal cortex, which is associated with inhibition of irrelevant speech (see Gisselgård et al., 2004).

A second possible explanation for the absence of an age by noise interaction is a differential development of global—or simple—
irrelevant inhibition, on the one hand, and selective inhibition, on the other hand. Where global inhibition requires the suppression of all background stimuli, which was the case in the current study, selective inhibition requires the suppression of only part of the background stimuli. In the domain of motor behavior, De Jong, Coles, and Logan (1995) showed that response inhibition in a stop-signal paradigm takes more effort if participants are instructed to interrupt their response after every stop signal than if they are instructed to do so only after a specific type of stop signals. Using a comparable paradigm, Van den Wildenbergh and Van der Molen (2004) found that these two types of inhibition develop differently in children; that is, selective inhibition develops more slowly than global inhibition. In an aging perspective, there could be a reversed development of these two types of inhibition; that is, selective inhibition could show a stronger age-related decline than global inhibition. If this idea is tested with the irrelevant-speech paradigm, then age effects should emerge if some part of the background speech is relevant and the other part is not. No such studies are reported in the context of the irrelevant speech paradigm, where the primary task is presented in the visual modality. On the other hand, disproportionate effects of irrelevant speech on performance of elderly individuals were found in studies of speech comprehension (e.g., Schneider, Daneman, Murphy, & Kwong See, 2000; Schneider, Daneman, & Pichora-Fuller, 2002; but see Murphy et al., 1999). These effects were attributed to age-related hearing loss rather than cognitive decline, however.

Finally, it is acknowledged that the irrelevant speech paradigm considerably differs from the more common inhibition paradigms in cognitive aging research, such as the negative priming paradigm. As Rouleau and Belleville (1996) already pointed out, most of these paradigms merely address the visual modality, whereas the irrelevant speech paradigm addresses both the visual and the auditory modality (but see Li et al., 2004). Furthermore, most inhibition paradigms require the participant to process both relevant and irrelevant information to perform the task successfully, whereas the irrelevant speech paradigm only requires the participants to process relevant information. These differences might indicate that inhibition of extraneous information is relatively easy in the irrelevant speech paradigm.

In sum, the current study is in line with prior research establishing the absence of a disproportionate irrelevant speech effect in aging. Rouleau and Belleville’s (1996) notion that in the elderly phonological and lexicosemantic inhibition are dissociated in that the former is
relatively intact while the latter is impaired, was not supported. Interestingly, the current and previous studies on the irrelevant speech effect in aging reject what Perfect and Maylor (2000) call the “dull hypothesis”: an age by complexity interaction in which age differences increase with the complexity of the task. Because this interaction is so common, Perfect and Maylor state that it hardly adds new insights into cognitive aging. In fact, they consider rejection of the dull hypothesis as the basis for developing theories of cognitive aging. From this perspective, it is ironic that researchers still strive to confirm the “dull hypothesis” in the irrelevant speech paradigm. What the current and previous results essentially show is that the inhibitory view of cognitive aging is far less universal than expected.

REFERENCES


