THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 1989 *41A* (1) 107–122

# Effects of Background Music on Phonological Short-term Memory

Pierre Salamé

*CNRS/INRS Laboratoire de Physiologie et Psychologie Environnementales, Strasbourg, France*

Alan Baddeley

*MRC Applied Psychology Unit, Cambridge, U.K.*



Immediate memory for visually presented verbal material is disrupted by concurrent speech, even when the speech is unattended and in a foreign language. Unattended noise does not produce a reliable decrement. These results have been interpreted in terms of a phonological short-term store that excludes non-speechlike sounds. The characteristics of this exclusion process were explored by studying the effects of music on the serial recall of sequences of nine digits presented visually. Experiment 1 compared the effects of unattended vocal or instrumental music with quiet and showed that both types of music disrupted STM performance, with vocal music being more disruptive than instrumental music. Experiment 2 attempted to replicate this result using more highly trained subjects. Vocal music caused significantly more disruption than instrumental music, which was not significantly worse than the silent control condition. Experiment 3 compared instrumental music with unattended speech and with noise modulated in amplitude, the degree of modulation being the same as in speech. The results showed that the noise condition did not differ from silence; both of these proved less disruptive than instrumental music, which was in turn less disruptive than the unattended speech condition. Theoretical interpretation of these results and their potential practical implications for the disruption of cognitive performance by background music are discussed.

There is abundant evidence that the immediate serial recall of visually presented verbal material is impaired when presentation is accompanied by speech, even when this speech is in a foreign language and the subject is

Requests for reprints should be sent to Pierre Salamé, CNRS/INRS Laboratoire de Physiologie et Psychologie Environnementales, 21 Rue Becquerel, 67087 Strasbourg Cedex, France.

Thanks are due to Othon Schneider and Charles Schnitzler for their technical assistance in preparing the recordings. We are also grateful to Ian Nimmo-Smith for statistical advice, and to Anne Cutler, Chris Darwin, and Roy Patterson for comments on an earlier draft.

instructed to ignore it (Colle & Welsh, 1976; Salamé & Baddeley, 1983). On the other hand, when unpatterned white or pink noise accompanies presentation of verbal material for immediate serial recall, the effects are very far from clear. Some studies have reported impairment of recall (e.g. Rabbitt, 1968; Wilkinson, 1975); others have found an improvement in memory performance (e.g. Daee & Wilding, 1977; Hockey & Hamilton, 1970; Wilding, Mohindra, & Breen-Lweis, 1982); yet others observe no effect of noise on retention (e.g. Davies & Jones, 1975; Murray, 1965; Salamé & Baddeley, 1983; 1987).

Why should speech have such a clear and unequivocal effect, in contrast to the highly equivocal data for noise? Salamé and Baddeley (1982) have interpreted the unattended speech effect within a working memory framework. They suggest that visually presented sequences of verbal material, such as digits, are remembered via the articulatory loop system. More specifically, the visual items are recoded through a process of subvocalization, a process that allows them to be stored within the short-term phonological store component of the articulatory loop. The unattended speech effect is assumed to occur because spoken material gains obligatory access to the phonological store, and hence is able to interfere with the use of that store for helping to retain the visually presented item. Evidence for this interpretation comes from studies in which subjects are prevented from recoding the visual items by means of articulatory suppression, the requirement to utter repeatedly some redundant word, a procedure that is assumed to prevent the subvocalization of the visually presented items. Under these conditions, the unattended speech effect is abolished (Baddeley, 1986; Salamé & Baddeley, 1982; 1983).

How would such an interpretation handle the observation that noise has such an equivocal effect on immediate memory? If one makes the plausible assumption that the effect of noise on the phonological store is analogous to the effect of noise on acoustic discrimination, then one would expect that although it might have less of a masking effect than speech, it would nevertheless cause some consistent impairment in memory performance, an impairment that is not reliably observed. Salamé and Baddeley (1987) suggest that one possible way of conceptualizing the situation is to assume that some kind of filter governs the access of sounds to the phonological store, keeping out noise but allowing in speech. Such a mechanism could of course be extremely useful in helping the organism perceive speech in a noisy environment. However, any filter must have certain characteristics that allow it to perform this function. This raises the issue as to what those characteristics might be. As will be discussed later, a filter is not the only way of conceptualizing these results. Nevertheless, the question of what makes speech speechlike remains important, whether one adopts a filter concept or opts for some alternative model.

One obvious difference between noise and speech is that speech is structured and patterned. If this is the crucial factor, then one might expect music, which is also a highly structured perceptual auditory input, to gain access to the phonological store and hence to disrupt performance. Another possibility is that the crucial feature is some aspect of the human voice, in which case instrumental music might be expected to function very much as noise, and not to disrupt performance on visually presented serial recall. This raises the further question of whether a human voice when singing would be characterized as more like music or more like speech. The experiments that follow explore these possibilities by studying the influence of vocal and instrumental background music on immediate memory for visually presented sequences of digits.

While the primary function of this study was theoretical, it also has implications for the applied question of what influence background music is likely to have on cognitive performance. It is not uncommon for work-places to have background music, and studying to a background of music is far from uncommon. Furthermore, with the development of cheap and effective personal stereo systems, it is increasingly common for music to accompany activities that would previously have been performed in silence. If it proves to be the case that music has a disrupting effect on immediate memory, the possibility is raised that it may impair performance on a much wider range of cognitive tasks.

The evidence for an influence of music on performance is at present rather sparse, typically being concerned with its effect on vigilance tasks in which it often tends to lead, if anything, to an improvement. For example Davies, Hockey, and Taylor (1969) found that a mixture of speech and music enhanced performance on a vigilance task, although only in extrovert subjects. In a subsequent study, Davies, Lang, and Shackleton (1973) compared the effect of background music and white noise on vigilance. They found better detection with a noise than a music background when an easy task was used, but the reverse for a more difficult task. A more recent review of the effects of background music is given by Fox (1983), who suggests that music may have a positive effect as an arouser, hence enhancing performance on dull and repetitive tasks, although no data on the effects of music on memory are reported. Should our own study reveal effects of music, then it would clearly point to the need to explore this issue more extensively, using rather more ecologically relevant tasks than the immediate recall of visually presented digits.

## EXPERIMENT 1

### Method

*Material and Design.* This study explores the effect of instrumental and vocal music on the immediate serial recall of sequences of nine visually presented digits. All subjects performed the task in three conditions that were counterbalanced using a $3 \times 3$ Latin square design. The conditions comprised:

S: a silent control condition with an ambient sound level of 37 dB(A);
V: vocal music played continuously at an average level of 75 dB(A) during presentation and recall;
I: instrumental music played continuously at an average level of 75 dB(A).



In this and subsequent experiments, the acoustic stimuli were presented by means of two Lansing loudspeakers. Each of the musical conditions involved four pieces, the vocal conditions comprising two male and two female singers, namely (a) *Débranche*, sung by France Gall; (b) *Le Vieux Chêne*, sung by Georges Brassens; (c) *Encore un Matin*, sung by Jean Jacques Goldmann; and (d) *Africa*, sung by Rose Laurens. The instrumental excerpts comprised (a) Ravel's *Bolero*; (b) Berlioz's *Rákóczi March*; (c) Alford's *Bridge over the River Kwai*; and (d) Offenbach's *French Cancan*. Half of the subjects experienced the excerpts in the order (a), (b), (c), (d) and the other half in the order (c), (d), (a), (b).

The sequences of nine digits were presented in the centre of the screen of a DAI micrcomputer at a rate of 1 per 750 msec. Each sequence was preceded by a 2-sec visual warning. Subjects were instructed to remain silent, but to group the digits mentally into threes and to write them down immediately after presentation in a response booklet. They were given 13 sec for recall and were requested to recall and write them in order of presentation. Subjects were instructed to concentrate on the memory task and to ignore the music.

All subjects began by performing the task in silence. They were given one block of nine practice sequences, and this was used to ensure that the three groups, given the three different orders of conditions, were approximately equal in competence. Subjects then proceeded to three sessions, each comprising four further blocks of nine sequences, one session in each of the three experimental conditions. A session lasted approximately 15 minutes. Each block of sequences lasted 3 min 30 sec and was separated from the next block by a gap of 6 sec. Each musical excerpt lasted for 3 min 36 sec and began 6 sec before the sequence of digits, so as to minimize any effects on memory of surprise when a new excerpt was introduced.

The subjects comprised 22 female and 2 male students from the Introductory Psychology Course of the University of Strasbourg. Their mean age was 19 years, and they were paid for participation. Half the subjects in each group were tested between 9.30 and 11.30 a.m., and half between 2.00 and 4.00 p.m.

## Results and Discussion

The overall performance of the three groups is shown in Figure 1. Analysis of variance showed no significant effect of the order in which conditions had been presented, $F(2, 12) = 1.21, p > 0.05$, no effect of the order in which the music was presented, ($F < 1$), and no effect of time of day ($F < 1$). The analysis did, however, show an effect of conditions, $F(2, 24) = 42.42$, $p < 0.001$, of serial position, $F(8, 96) = 41.81$, $p < 0.001$, and a significant Condition × Serial position interaction, $F(16, 192) = 4.83, p < 0.05$. None of the other interactions reached significance. A Newman-Keuls tests indicated that vocal music led to significantly more errors than did instrumental music, which, in turn, produced significantly more errors than the silent control condition, $p < 0.01$ in both cases.

To what extent are the results we obtained from an analysis across subjects supported when the analysis is carried out across musical selections rather

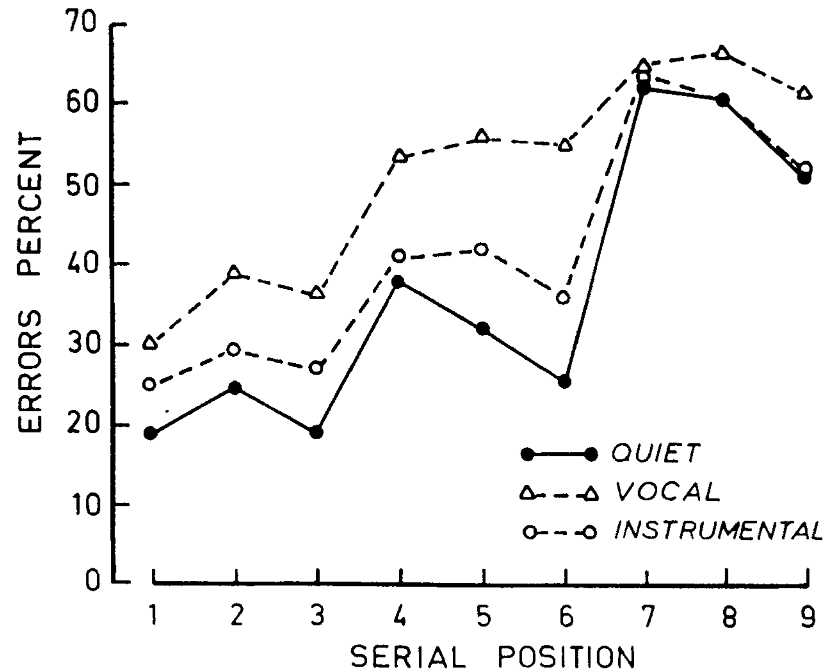


FIG. 1. The influence of acoustic background on the immediate serial recall of digits: Results of Experiment 1.

than subjects? Table 1 shows the mean percentage error rate obtained for each of the musical selections, and for the four periods during the silent control that would be equivalent in the counterbalanced design to the four passages in each of the music conditions. Table 1 suggests first of all that the results obtained across subjects also hold across passages; there is no overlap between the mean error rates in the silent control and in the instrumental music condition, nor between the instrumental and vocal error rates.

Analysis of variance across passages again showed a significant effect of conditions, $F(2, 9) = 77.18$, $p < 0.0001$. Combining analyses across subjects and conditions yielded $Min\ F'(2, 32) = 27.38$, $p < 0.0001$. Within the orchestral and vocal conditions, no significant differences were observed; this analysis was, of course, performed across subjects, $F(3, 36) = 1.49$, $p > 0.05$. This lack of difference is striking when one bears in mind the fact that the individual pieces varied considerably. For example, the vocal pieces ranged from a male singer, Brassens, with a complex narrative song accompanied by a single guitar to the song by France Gall, which contains much less dense narrative and is accompanied by a full orchestra.

The interaction with serial position appears to stem from the fact that the difference between conditions is greatest for the middle serial positions and least for the last three positions, where error rate is extremely high. Previous studies (e.g., Salamé & Baddeley, 1982) suggest that the unattended speech effect often interacts with serial position, but that the nature of the interaction is very inconsistent from one experiment to the next, in contrast to the very robust main effect of unattended speech. Hence, rather than attempt a theoretical interpretation of this interaction here, we will delay discussion until two further studies have been described.

We obtained no effect of time of day, a result that is broadly consistent with the findings of Loeb, Baker, and Holding (1983) and Baker, Holding, and Loeb (1984) who observed time-of-day effects for males, but not for females; our subjects were, of course, almost all females. Finally, the lack of

TABLE 1
Experiment 1: Mean Percentage Error Rate as a Function of Musical Background

| *Vocal* | | | *Orchestral* | | | *Silent Control* | |
|---|---|---|---|---|---|---|---|
| (a) | France Gall | 48.6 | (a) | Ravel | 40.2 | (a) | 35.8 |
| (b) | Georges Brassens | 53.1 | (b) | Berlioz | 38.1 | (b) | 37.2 |
| (c) | Jean Jacques Goldmann | 52.2 | (c) | Alford | 42.7 | (c) | 36.0 |
| (d) | Rose Laurens | 49.1 | (d) | Offenbach | 44.7 | (d) | 37.4 |
| | *Mean* | 50.8 | | | 41.4 | | 36.6 |

any interaction between conditions and order of presentation suggests that our results are unlikely to have been distorted by the presence of asymmetrical transfer effects (Poulton & Edwards, 1979).

## EXPERIMENT 2

Whereas the results of the first study appeared to be clear-cut, as the result is potentially an important one, we felt that it should be replicated before publication. Furthermore, we subsequently noted that the vocal passages were modern and in French, whereas the instrumental passages tended to be nineteenth century. Hence, it could be argued that the vocal music would be of greater appeal to the subjects, and, as the lyrics were in French, could be understood by the subjects and might be more distracting. We therefore decided to replicate the study using classical operatic arias in a foreign language in the vocal condition, and popular modern instrumental music.

### Method

*Material and Design.* Subjects performed the same task as in Experiment 1, but differed in that they had previously performed the memory task for an hour in a separate experiment; they would therefore be considerably more practised than the subjects in Experiment 1. Sound levels and order of conditions were identical with Experiment 1. The four arias comprised (a) *O zittre nicht mein lieber Sohn*, from Mozart's "Magic Flute"; (b) *La Calumnia è un venticello*, from Rossini's "Barber of Seville"; (c) *Das Wandern*, from Schubert's "Die Schöne Müllerin"; and (d) *Porgi amor qualche ristoro*, from Mozart's "Marriage of Figaro". (a) and (d) were female arias, (b) and (c) were sung by male singers. The instrumental pieces comprised (a) *Apache*, by The Shadows, (d) Mike Oldfield's *Tubular Bells*, (c) Duke Ellington's *Satin Doll*, and (d) *The Lebanon*, by Human League. Each passage lasted 3 min 35 sec and was presented continuously, in synchrony with each block of nine digits. Again two orders were used, (a), (b), (c), (d) and (c), (d), (a), (b).

The subjects were 24 male volunteers aged between 25 and 40 years, who had answered an advertisement inviting them to participate in an experiment on the affects of alcohol, which ran in parallel with the present study. Half were high-consumption social drinkers and half low. This variable was counterbalanced across order of conditions. All subjects were sober during testing, which occurred between 2.00 and 4.00 p.m. Again, subjects were instructed to ignore the music.

### Results and Discussion

The results of Experiment 2 are shown in Figure 2, from which it is clear that

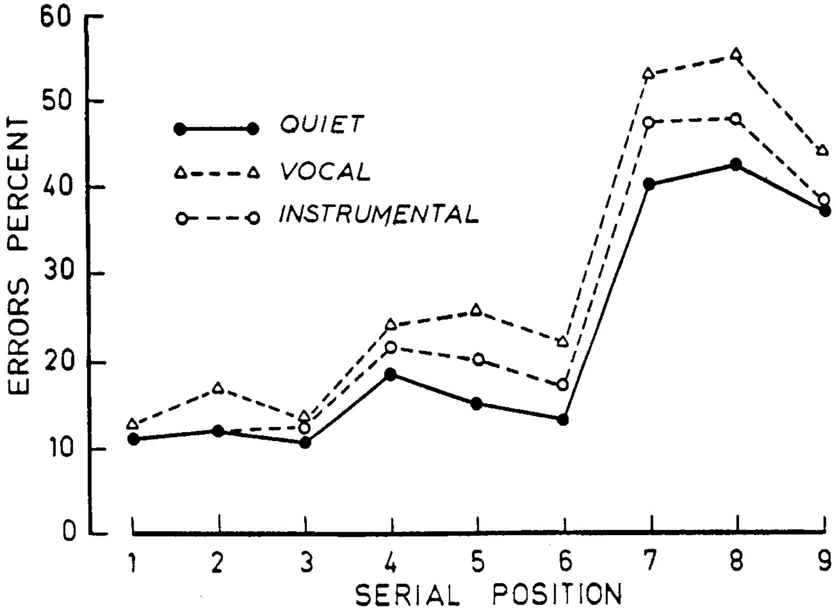


FIG. 2. Results of Experiment 2: Immediate serial recall of digits as a function of acoustic background.



overall level of performance was somewhat higher than observed in Experiment 1, possibly because of the hour of practice preceding the study. No differences were observed between high and low social drinkers, nor did this interact with any other variable. Analysis of variance did, however, show significant effects of conditions, $F(2, 36) = 9.84$, $p < 0.001$, and of serial position, $F(8, 144) = 32.21$, $p < 0.001$, together with a significant Conditions × Serial position interaction, $F(16, 288) = 2.12$, $p < 0.01$. A Newman-Keuls test indicated a significant difference between the quiet and vocal conditions, $p < 0.01$, and between the instrumental and vocal, $p < 0.05$, whereas the instrumental–quiet difference was not statistically significant.

Table 2 shows the results obtained across passages. Analysis of variance indicated a significant difference between conditions, $F(2, 9) = 9.94$, $p < 0.01$. Combining the analysis by subjects and by passages yielded *Min* $F'(2, 29) = 4.94$, $p < 0.05$.

These results, then, are broadly consistent with those of the first study. The vocal music condition leads to the highest error rate, followed by the instrumental music and the silent control. However, the difference between the instrumental and the silent condition in this case does not reach significance, possibly because the range of errors is rather smaller in this study than in the previous one. Indeed, the error rate for the first six serial

TABLE 2
Experiment 2: Mean Percentage Error-rate as a Function of Musical Background

| *Vocal* | | | *Instrumental* | | | *Silent Control* | |
|---|---|---|---|---|---|---|---|
| (a) | Mozart (♀) | 29.3 | (a) | The Shadows | 24.4 | (a) | 24.2 |
| (b) | Rossini (♂) | 29.3 | (a) | Mike Oldfield | 25.5 | (b) | 22.9 |
| (c) | Schubert (♂) | 34.6 | (c) | Duke Ellington | 26.5 | (c) | 21.4 |
| (d) | Mozart (♀) | 25.4 | (d) | Human League | 24.7 | (d) | 21.3 |
| | *Mean* | 29.7 | | | 25.3 | | 22.2 |

(♂) Male singer (♀) Female singer



positions is relatively low for all three conditions, with the last three serial positions showing the greatest separation, in clear contrast to Experiment 1. This suggests that the interaction with serial position observed in the two present experiments may be produced by ceiling and floor effects that attenuate differences between conditions.

The second study agreed with Experiment 1 in indicating very similar results for the four passages within the two music conditions. Furthermore, by showing a clear effect on performance of arias sung in a foreign language, they allow us to rule out the suggestion that the effect of vocal music in Experiment 1 may have been due to the fact that the vocal items were in the native language of the subjects, as in this study none of the arias was sung in French.

In one respect, however, the position is still somewhat unsatisfactory, because we now have one study indicating that instrumental music does impair performance, and one that fails to observe a significant effect. For that reason, we included an instrumental music condition in a subsequent experiment, which was primarily concerned with aspects of unattended speech. A central aim of this experiment was to test the hypothesis that fluctuations in sound intensity were a crucial factor in the unattended speech and music effects. This was tested by using noise that varied in amplitude in exactly the same way as an equivalent passage of unattended speech.

## EXPERIMENT 3

### Method

*Design and Procedure.* This experiment involved four conditions, a quiet control condition at a level of 37 dB(A) and three experimental conditions, each at a level of 75 dB(A). In each condition, subjects attempted to recall 27 sequences of nine visually presented digits, and to ignore any concurrent sounds. One of the conditions involved a continuous reading of Arabic prose (taken from "The Prophet" by G. K. Gibran), a second comprised three continuous passages of modern instrumental music (passages (a), (b) and (c) from Experiment 2). The third condition comprised modulated "pink" noise; this differs from white noise in that sound pressure is held constant across all the constituent octaves. The amplitude of the noise was driven by the amplitude of the Arabic speech played through a BK 3347 dB(A) filter with a time constant of 200 msec and a coefficient of 50 msec. The Arabic speech was used to drive a pink noise generator, which, in turn, fed a voltage-controlled amplifier. In this condition then, overall intensity of the noise mirrored that of speech, as is shown in Figure 3. Consequently, if the critical feature of speech for disrupting memory is its particular pattern of fluctuation in loudness, then this condition should disrupt performance in the same way as the unattended speech on which it was based.

Subjects performed a practice session in which they attempted to recall 18 sequences, each comprising nine digits, before going on to the main experiment in which they were tested in each of the four conditions, the order being determined by a $4 \times 4$ Latin square. Performance on the last nine practice sequences was used to assign the subjects to four matched groups. The subjects were 24 female psychology students from the University of Strasbourg, aged between 19 and 22 years. None had previously taken part in a memory experiment, all were paid for participation and were tested in the afternoon.



## Results

The results of this study are shown in Figure 4. Analysis of variance showed a significant effect of conditions, $F(3, 60) = 42.53$, $p < 0.001$, and of serial position, $F(8, 160) = 21.24$, $p < 0.001$. Serial position again interacted with condition, $F(24, 48) = 3.66$, $p < 0.01$. The only other significant interaction was that between groups and conditions, $F(9, 60) = 3.42$, $p < 0.01$. A Newman-Keuls test performed on number of errors per condition showed that the noise and silent conditions did not differ, both being significantly better than instrumental music, $p < 0.01$, which in turn was better than the speech condition, $p < 0.01$.

Although the overall results of this study are clear-cut, two points require further comment. The first of these concerns the interaction with serial position, where again the effects appear to be rather less strong over the last three serial positions than over the first six, a result analogous to that obtained in Experiment 1 and inconsistent with the results of Experiment 2. We suspect that this simply represents a tendency for the unattended speech

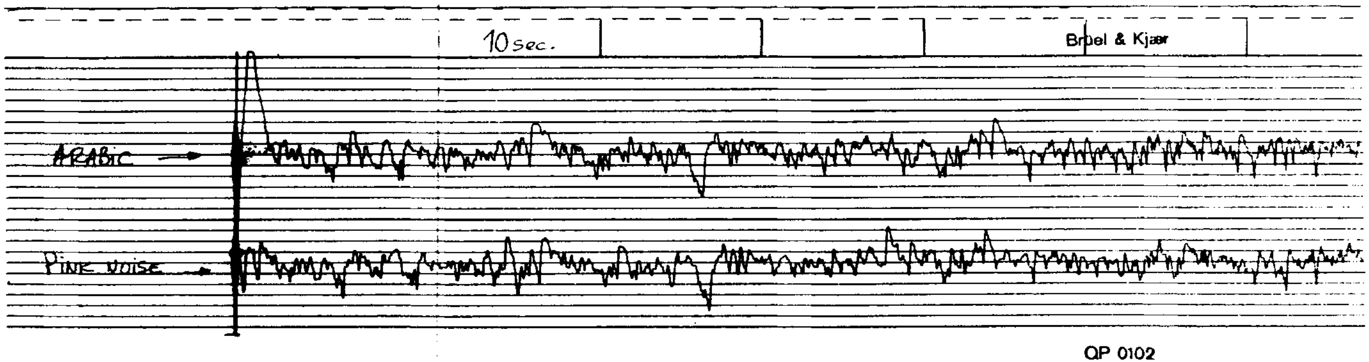


FIG. 3. Intensity envelope for the Arabic speech and the amplitude modulated noise used in Experiment 3.

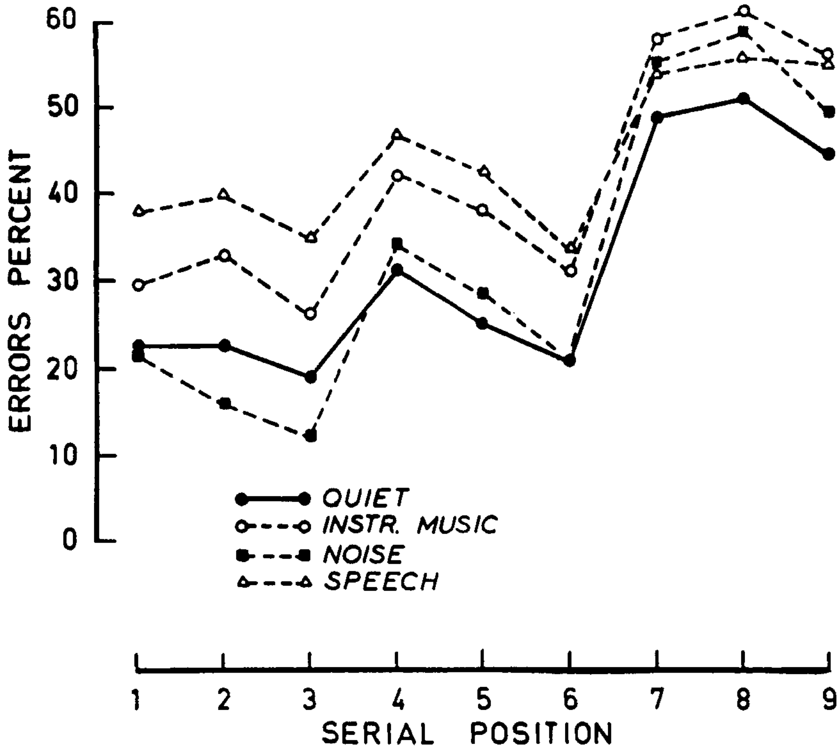


FIG. 4. Mean percentage errors in recalling digits against a background of quiet, modulated noise, instrumental music, and speech.

effect to be less pronounced when performance is at a very low level, possibly because the subject ceases to use a phonological code, relying on other sources of information. Evidence for this comes from an earlier study (Salamé & Baddeley, 1986) in which the effects of unattended speech and phonological similarity were studied in a task involving the serial recall of sequences ranging from five to eight letters. While both effects were clear for shorter sequences, performance at length eight tended to show little effect of unattended speech when combined with phonological similarity.

A slightly more worrying result is the Groups × Conditions interaction, which suggests the possibility of a distortion of results by asymmetric transfer effects (Poulton & Edwards, 1979). In order to explore this possibility further, we performed a further analysis based only on the first condition tested, and hence equivalent to a separate groups design. In order to increase the power of the analysis and minimize effects of individual differences, we used pretest practice scores, all of which took place in silence, as a co-variate. Planned comparison between the critical conditions of music and silence indicated a statistically significant difference, $t = 2.66$, $df = 19$; $p < 0.01$, one-tailed. Inspection of the data suggested that the Groups × Conditions interaction was largely attributable to the fact that on the first trial block the music condition led to a non-significantly greater number of errors than did the speech condition, whereas overall pattern indicated a higher error rate for the speech than the music condition.

## GENERAL DISCUSSION

Experiments 1 and 2 showed clear effects of vocal music on immediate memory for visually presented sequences. Experiments 1 and 3 showed that performance was also impaired by instrumental music, with Experiments 1 and 2 indicating that this effect is significantly less than that found with vocal music. At an empirical level, then, our results are simpler than we had anticipated. We included a wide range of both vocal and instrumental music, in the expectation that if music had any effect, then differences between excerpts of vocal or instrumental music would occur, and would offer a clue to the critical underlying acoustic variables. In fact, we found no solid evidence for differences between the various vocal or instrumental excerpts, suggesting that speculation on the critical features would be unwise. It seems intuitively likely that differences will occur if sufficiently diverse samples are used, but we suspect that a more fruitful approach may be to generate sounds that vary systematically in specified dimensions such as pitch, timbre, prosody, and rhythm, rather than to rely on precomposed music.

What are the practical implications of our findings? They suggest that background music can certainly disrupt immediate verbal memory, particularly in the case of vocal music. This suggests the possibility that a similar disruption may occur for other more ecologically important cognitive tasks that also use the phonological store. Possible tasks include reading, counting, calculating, and reasoning (Baddeley, 1986). In view of the prevalence of background music in both working and studying environments, this issue should clearly be explored more widely.

What do our results imply theoretically? We suggested in the introduction that our earlier results were consistent with the concept of a phonological store that could be used to supplement the recall of visually presented verbal material. Spoken material was assumed to gain obligatory access to this store, where it is able to interfere with the retention of the recoded visual items. The fact that in our earlier studies, and in the third experiment of the present series, white noise does not impair performance suggests that it does not gain access to the store. The present series of experiments was concerned with beginning to outline those characteristics of speech that allow it to be registered in the store. Our present results are consistent with the previously described hypothesis of a filter that allows in speech and cuts out noise, but only if one assumes that the filter's operation is not all-or-none, as instrumental music appears to have some effect, but not as great as that involving vocalization.

How could such a filter work? Presumably it would need to govern access to the phonological store on the basis of a range of acoustic cues that are present in speech, but not in noise. There are, naturally, many candidates for this, including the periodicity that occurs in vowel sounds, and of course in music, the fluctuations in loudness, in pitch, or indeed some combination of these.

The fact that instrumental music is intermediate in its disruptive effect between speech and silence, however, implies the need to assume something more than an all-for-none filter. The differential degree of disruption could be explained in two possible ways. The first is to assume that music contains a sub-sample of features that mimic that of speech. When such features occur, they will gain access and disrupt performance, but they will be intermittent and hence less disruptive than speech itself. If this were so, one might perhaps expect different types of instrumental music to be differentially disruptive. As we saw, no evidence of this was found.

A more attractive hypothesis might be to assume that the unattended speech effect depends on two separable functions, the first being concerned with whether the unattended material passes the filter, the second with the masking effects achieved by whatever material gains access to the phonological store. The lack of an effect of noise on memory would then be accounted for in terms of the first assumption, namely that noise does not have the necessary characteristics to pass the peripheral filter. The difference in effect between vocal and instrumental music, on the other hand, would be explained in terms of their differential capacity to disrupt the phonological store. Vocal music, having more acoustic features in common with subvocal speech, would be more likely to disrupt it.

Some evidence for phonological factors in the disruption of memory comes from an earlier study (Salamé & Baddeley, 1982), in which it was shown that disruption of memory for visually presented digits was a function of the phonological similarity between the digits and the unattended material. Performance was less disrupted by unattended disyllabic words selected so as to be phonologically dissimilar to digits than it was by words made up from the same phonemes as the digits (e.g. *tun*, *woo* instead of *one*, *two*). This two-stage hypothesis can best be tested by using spoken material, but systematically studying the relationship between the material to-be-remembered and the unattended auditory material.

Our experiments were based on the concept of an acoustic filter and were aimed at investigating its broad characteristics. However, as C. J. Darwin (personal communication) has pointed out, our results could also be conceptualized in terms of a mechanism set up to detect speech rather than to filter out noise. The distinction between a negative filter and a positive detector might be clarified by using two physical analogies. The process of filtering involves allowing some things through but not others, as in the case of a coffee filter where the coffee grounds are held by the filter, while the liquid passes through. One can contrast this with a detection system such as a photographic plate, where it is probably more useful to think of the system as being sensitive to certain light wave lengths, rather than as a system for filtering out other forms of energy such as heat and sound. We suspect that these two analogies are both applicable to our results; our experiments were not, of course, designed to decide between them. The two do, however, differ in their implications for the overall architecture of the system, and for that reason the detector analogy is worth exploring in further depth.

How might such a phonemic detector explain our results? Underlying this approach is the assumption that the short-term storage component of the articulatory loop system is not a general acoustic store protected by a filter, but is rather a detection and storage system that has evolved for processing speech-like material. On this interpretation, white noise would have little effect, as it would not have the characteristics that would register in such a store, whereas vocal and, to a lesser extent, instrumental music would.

There are a number of attractive features to this view, one being that of comparative simplicity. One does not need to assume separate processes for explaining the absence of an effect of noise and for the presence of a moderate effect of instrumental music—both can be assumed to reflect the extent to which the respective sounds are capable of registering within the phonological storage system. This economy may not be quite as great as it seems at first, however, because the fact that we can hear and remember sounds that are very unlike speech means that there must be some—presumably additional—form of acoustic storage system capable of dealing with such material.

The second reason for the attractiveness of the phonological detector hypothesis is that it implies clear and strong assumptions about the relationship between phonological storage in STM paradigms and the processes of speech perception, as the characteristics of the positive detection system are presumably central to the normal process of detecting and perceiving spoken sounds. This, in turn, suggests the fruitfulness of seeking a common model for both phonological perception and storage. Patterson's (1986) spiral processing model offers an interesting candidate for the possible detection process, suggesting that a more detailed modelling of our results may also require both filtering and detection processes.

In conclusion, whether the better conceptualization of the unattended speech effect is in terms of a filter or a detector, it is becoming increasingly clear that understanding the memory phenomena will depend on understanding the basic processes of speech perception.

## REFERENCES

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baker, M. A., Holding, D. H., & Loeb, M. (1984). Noise, sex and time of day effects in a mathematics task. *Ergonomics, 27*, 67–80.

Colle, H. A. & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15*, 17–32.

Daee, S. & Wilding, J. M. (1977). Effects of high intensity noise on short-term memory for position in a list and sequence. *British Journal of Psychology, 68*, 335–349.

Davies, D. R., Hockey, G. R. J., & Taylor, A. (1969). Varied auditory stimulation, temperament differences and vigilance performance. *British Journal of Psychology, 60*, 453–457.

Davies, D. R. & Jones, D. M. (1975). The effects of noise and incentive in short-term memory. *British Journal of Psychology, 66*, 61–68.

Davies, D. R., Lang, L., & Shackleton, V. J. (1973). The effects of music and task difficulty on performance at a visual vigilance task. *British Journal of Psychology, 64*, 383–389.

Fox, J. G. (1983). Industrial music. In D. Oborne & M. Gruneberg (Eds.), *The physical environment at work*. New York: Wiley.

Hockey, G. R. J., & Hamilton, P. (1970). Arousal and information in short-term memory. *Nature, 226*, 866–867.

Loeb, M., Baker, M. A., & Holding, D. H. (1983). Sex, arousal and fatigue determining noise effects and aftereffects. In G. Rossi (Ed.), *Proceedings of the 4th international congress on noise as a public health problem, Vol. 2* (pp. 787–796). Milan: Centro Ricerche e Studi Amplifon.

Murray, D. J. (1965). The effect of white noise upon the recall of vocalized lists. *Canadian Journal of Psychology, 19*, 333–345.

Patterson, R. D. (1986). Spiral detection of periodicity and the spiral form of musical scales. *Psychology of Music, 14*, 44–61.

Poulton, E. C. & Edwards, R. S. (1979). Asymmetrical transfer in within-subject experiments on stress interactions. *Ergonomics, 22*, 945–961.

Rabbitt, P. M. A. (1968). Channel capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology, 20*, 241–248.

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

Salamé, P. & Baddeley, A. D. (1983). Differential effects of noise and speech on short-term memory. In G. Rossi (Ed.), *Proceedings of the 4th international congress on noise as a public health problem, Vol. 2* (pp. 751–758). Milan: Centro Ricerche e Studi Amplifon.

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

Salamé, P. & Baddeley, A. D. (1987). Noise, unattended speech and short-term memory. *Ergonomics, 30*, 1185–1193.

Wilding, J., Mohindra, N., & Breen-Lewis, K. (1982). Noise effects in free recall with different orienting tasks. *British Journal of Psychology, 73*, 479–486.

Wilkinson, R. T. (1975). Noise, incentive and prolonged work: Effects on short-term memory. Paper presented at the annual meeting of the American Psychological Association, Chicago.

*Revised manuscript received 18 March 1988*