SPECIAL ISSUE: ORIGINAL ARTICLE

FROM DICHOTIC LISTENING TO THE IRRELEVANT SOUND EFFECT: A BEHAVIOURAL AND NEUROIMAGING ANALYSIS OF THE PROCESSING OF UNATTENDED SPEECH

C. Philip Beaman[1], Andrew M. Bridges[2] and Sophie K. Scott[3]

([1]Department of Psychology, University of Reading, Reading, UK; [2]School of Applied Sciences, University of Wolverhampton, Wolverhampton, UK; [3]Institute of Cognitive Neuroscience, University College London, London, UK)

Abstract

The assumption that ignoring irrelevant sound in a serial recall situation is identical to ignoring a non-target channel in dichotic listening is challenged. Dichotic listening is open to moderating effects of working memory capacity (Conway et al., 2001) whereas irrelevant sound effects (ISE) are not (Beaman, 2004). A right ear processing bias is apparent in dichotic listening, whereas the bias is to the left ear in the ISE (Hadlington et al., 2004). Positron emission tomography (PET) imaging data (Scott et al., 2004, submitted) show bilateral activation of the superior temporal gyrus (STG) in the presence of intelligible, but ignored, background speech and right hemisphere activation of the STG in the presence of unintelligible background speech. It is suggested that the right STG may be involved in the ISE and a particularly strong left ear effect might occur because of the contralateral connections in audition. It is further suggested that left STG activity is associated with dichotic listening effects and may be influenced by working memory span capacity. The relationship of this functional and neuroanatomical model to known neural correlates of working memory is considered.

Key words: auditory attention, ignored speech, informational-masking

Introduction

As acknowledged by various authors (e.g., Baddeley and Weiskrantz, 1993; Cowan, 1995) the seminal work of Broadbent (1958) on selective auditory attention has proven highly influential. This influence extends to numerous later studies, not only of auditory attention *per se,* but also of many related issues concerning the interaction between attention and primary memory. One stream of such research has been the use of dichotic listening as a technique for assessing various aspects of auditory cognition. Another stream has encompassed investigations into the disruptive effects of supposedly unattended sound on cognition, the so-called "irrelevant sound effect" (ISE) (Beaman and Jones, 1997). Thus, it can be argued that investigations of dichotic listening and of auditory distraction by irrelevant sound stem from the same original research questions. Despite this shared root, there have been surprisingly few attempts to investigate how widely applicable the findings in each paradigm might be by exporting predictions derived from one situation to the other.

The ISE can be characterized by the observation that, when attempting an immediate recall task, the presence of task-irrelevant background speech generally increases the error rate on the primary task by between 13% (Neath et al., 2003) and 50% (Ellermeier and Zimmer, 1997). Researchers interested in the ISE have frequently assumed that there is a relationship between the ISE and dichotic listening. For example, Beaman and Jones (1997) suggested that, in both paradigms, the main aim is to understand the extent of analysis of an unattended message. In the ISE paradigm, this is achieved by examining how the characteristics of the to-be-ignored irrelevant sound influence performance on the primary memory task. The dependent variable in this case is the extent of disruption to the immediate memory task caused by the presence of irrelevant sound.

An initial finding (Jones and Macken, 1993) was that irrelevant non-speech sound can be as effective as irrelevant speech in disrupting performance on a primary task (usually, but not uniquely, immediate serial recall; Beaman and Jones, 1997, 1998; LeCompte, 1994). Subsequent studies, more closely examining the nature of the processing undergone by the irrelevant sound have determined that irrelevant sounds can be quite effectively "streamed", or perceptually grouped, either by pitch or by location. Such grouping can reduce the size of the ISE (Jones and Macken, 1995; Macken et al., 2003; Nicholls and Jones, 2002). Conversely, multiple streams originating from the same spatial location produce a perceptual "babble" of noise that is less disruptive to serial recall than a single stream that alternates abruptly in one or more dimensions [e.g., pitch (Jones and Macken, 1993) or lexical identity (Bridges and Jones, 1996)]. Other findings from the irrelevant sound paradigm suggest that unattended speech is analysed for meaning but that this affects immediate recall only when the primary recall task also requires extensive semantic processing (Beaman, 2004; Buchner et al., 1996; Jones et al., 1992; Neely and LeCompte, 1999; Oswald et al., 2000; but see Buchner et al., 2004; Buchner and Erdfelder, 2005).

One intuition regarding the irrelevant sound is that there should be some relationship between immediate memory capacity and the effects of irrelevant sound, where capacity is measured either by span tasks (Beaman, 2004; Elliott and Cowan, 2005; Neath and Surprenant, 2004) or by number correct in a quiet control condition (Ellermeier and Zimmer, 1997; Neath et al., 2003). A key tenet of several influential models of immediate memory is that they assume some form of restricted capacity – either within a phonological store (Baddeley, 1986) or upon an episodic surface (Jones et al., 1996). Within these models, the ISE is neatly accounted for by the suggestion that unwanted irrelevant sounds take up "space" that would otherwise be available to store relevant, to-be-recalled material. If this view is taken, then those participants with more storage space available should be less susceptible to the ISE and there should be a negative correlation between short-term memory capabilities and the susceptibility of participants to irrelevant sound.

On the basis of this intuition, a strong relationship of some kind between immediate memory capacity and the ISE seems plausible. However, studies across several laboratories and using several different techniques have failed to find such a relationship (Ellermeier and Zimmer, 1997; Elliott and Cowan, 2005; Neath et al., 2003). Multiple measures were employed; including both absolute and percentage ISE measures, together with different means of measuring span (see Ellermeier and Zimmer, 1997; Elliott and Cowan, 2005; Neath et al., 2003; Neath and Surprenant, 2004, for full details of each set of studies). As in other fields of investigation, however, it is possible that working memory tasks more complex than the traditional short-term memory span technique may prove more informative (Engle et al., 1999). Complex working memory span tasks have, for example, proved useful when applied to dichotic listening situations (Conway et al., 2001).

The Role of Working Memory in the Ise and the Cocktail Party Phenomenon

Dichotic listening studies of the extent and nature of processing unattended speech have a longer history than ISE studies. Many dichotic listening studies used the "shadowing" technique, whereby participants were required to repeat the to-be-attended message as it was presented to one ear while the supposedly irrelevant, to-be-ignored message was presented to the other ear. The early studies of this kind by Broadbent (1958) and Cherry (1953) are well-known and their findings broadly accepted. These include the failure to detect when the ignored speech is in a foreign language or played backwards (Cherry, 1953) foreshadow the finding that foreign or reversed speech disrupts serial recall to the same extent as forward speech (Jones et al., 1990). Further findings include the observation that pitch changes in the to-be-ignored speech (e.g., a change from male to female speaker or the insertion of a pure tone) are frequently noticed by the participants, but that one's own name is only noticed on the to-be-ignored message approximately one third of the time (Moray, 1959). Many of these findings have been confirmed in more recent tests (e.g., Wood and Cowan, 1995).

The idea that both the irrelevant sound and the dichotic listening paradigms are primarily concerned with the analysis of unattended sound (Beaman and Jones, 1997; Macken et al., 2003) potentially makes the findings from one research program, such as an effect of working memory span, relevant to understanding performance in the other paradigm. Making use of the supposed commonality in the object of study is a strategy adopted in recent studies by Beaman (2004) and Hadlington et al. (2004). Both Beaman (2004) and Hadlington et al. (2004) imported findings from the literature on dichotic listening, and in particular the "cocktail party effect" (Cherry, 1953) in an attempt to further understand the ISE.

The study by Beaman (2004) was based upon a previous report by Conway et al. (2001). This earlier report demonstrated that participants with high working memory capacity, as measured by the operation span (OSPAN) task, were less likely than participants with low OSPAN to notice their own name if presented to the unattended ear in a dichotic listening task. The OSPAN task is a variant of a family of complex working memory tasks first introduced by Daneman and Carpenter (1980). These tasks attempt to measure the ability to successfully share resources or alternate between processing and storage demands. In the OSPAN task, introduced by Turner and Engle (1989), a mathematical operation is presented to the participant (e.g., 'Is 5/2 = 3?'). A word is also presented to the participant concurrently with the operation. Participants are required to read aloud both operation and word and also to verify whether the operation is correct or not (usually by clicking on a button marked either yes or no). Once several such operations and words have been presented to the participant, the participant is cued to recall all the words they saw in the correct serial order.

One interpretation of the OSPAN and related tasks is that they measure a form of executive control of attention (Engle et al., 1999) and this interpretation was used by Conway et al. (2001) to explain their own results on the dichotic listening task (see also Engle, 2001). Specifically, this interpretation states that working memory capacity is "responsible for maintaining activation to relevant information and *suppressing distracting information*" (Conway and Engle, 1994, emphasis added). Individual differences in the appearance of the cocktail party effect presumably occur because of the different levels of control individuals are able to exert over their own attentional focus. This situation maps onto Cowan's (1995) "embedded processes" model of working memory fairly neatly, wherein irrelevant sound diverts attentional resources away from the to-be-recalled material. The Conway et al. (2001) dichotic listening data can be explained by assuming that OSPAN reflects a level of attentional control – an ability to prevent attentional resources from being diverted towards the irrelevant sound. If such attentional control abilities are unrelated to short-term memory storage capacity this also accounts for the unexpected absence of a strong correlation between short-term memory span and the ISE.

OSPAN, according to this argument, is sensitive to this attentional control capability. High-span participants are better able to focus their attention on the immediate task in hand and therefore, by extension, are less likely to notice their own name in a to-be-ignored message (Conway et al., 2001; Engle, 2001). Since it has been established by previous research that reliable individual differences exist between individuals in their susceptibility to the ISE (Ellermeier and Zimmer, 1997) it is reasonable to suggest that OSPAN might be linked to such differences in the same way as it is linked to individual differences in the cocktail party phenomenon. The underlying assumption here is that sensitivity to the content of the unattended message in Conway et al.'s (2001) dichotic listening task might map onto similar sensitivity when participants are attempting to ignore task-irrelevant speech whilst simultaneously carrying out an immediate memory task.

In addition to reasoning from the precedent set by the dichotic listening technique, OSPAN is an appropriate measure for testing individual differences in the ISE on the grounds that complex working memory span tasks such as OSPAN frequently predict susceptibility to distraction of various kinds where short-term memory span tasks do not (see Engle, 2001 and Engle et al., 1999 for reviews). For example, working memory span tasks such as OSPAN have been shown to predict susceptibility to proactive interference (Rosen and Engle, 1998) negative priming (Engle et al., 1995) and Stroop interference (Kane and Engle, 2000) as well as Conway et al.'s (2001) results with dichotic listening. ISEs also show evidence of negative priming (Driver and Baylis, 1993; Hughes and Jones, 2003, 2005) and there is some evidence also for a "cross-modal" Stroop effect with irrelevant speech (Cowan and Barron, 1987; Driver and Baylis, 1993; Elliott et al., 1998). This further suggests that OSPAN might provide an appropriate basis for characterizing individual differences in susceptibility to irrelevant sound.

Contrary to initial predictions, however, and despite several attempts, Beaman (2004) was unable to find an effect of OSPAN on the magnitude of the ISE, a result that has subsequently been independently replicated by Neath and Surprenant (2004). These failures, across two separate studies and three experiments with a total n of 218, contrast markedly with the ease with which a relationship between dichotic listening and OSPAN appeared for Conway et al. (2001; one experiment with n = 80). It is concluded that if there is any relationship between the faculty for controlled attention represented by working memory span and individual differences in susceptibility to the ISE, it is too weak to become apparent across studies by Beaman (2004) and Neath and Surprenant (2004).

This failure to observe a span/ISE relationship calls into question the assumption by Beaman and Jones (1997) that the dichotic listening technique and the ISE test the same mechanisms. The conclusion from the study by Conway et al. (2001) is that dichotic listening is related to working memory span (participants with high span more easily block out the irrelevant message). The conclusion from Beaman (2004) and elsewhere (Elliott and Cowan, 2005; Neath and Surprenant, 2004) is that ISE is not related to working memory span. Our first inference, therefore, is that the effects observed in the two paradigms involve different mechanisms. Importantly, however, one experiment by Beaman (2004) did show an effect of OSPAN. This effect was upon the number of intrusions from the irrelevant speech stream mistakenly recalled as present in the to-be-attended information visually-presented for free recall. These data are reproduced in Table I. The to-be-attended lists comprised of low-typicality members of a particular category and the to-be-ignored lists in the related condition were high-typicality members of the same category or, in the unrelated condition, high-typicality members of a different category.

The table shows that low-span participants were more likely than high-span participants to mistakenly recall irrelevant speech items when the irrelevant speech was semantically related (came from the same category) as the to-be-recalled items. One important factor is that this experiment incorporated a control for guessing. The number of intrusions observed in the "quiet" condition

TABLE I

*Mean number of intrusion errors made by high and low span participants in each irrelevant speech condition. Data taken from Beaman (2004)**

	Quiet	Unrelated speech	Related speech
High span	2.17 (.64)	4.39 (1.19)	5.33 (1.22)
Low span	2 (.51)	5.95 (.98)	8.74 (1.22)

*Standard errors are given in parentheses.

represents the likelihood of a categorically-related item intruding in the absence of any irrelevant speech. The number of intrusions observed in the "unrelated" condition represents the likelihood of a categorically-related item intruding in the presence of unrelated speech. Thus, the quiet condition provides a conservative measure of guessing, and the unrelated condition provides a less conservative measure of guessing when performance is generally depressed and acceptance thresholds might be strategically lowered. The results, therefore, reflect the availability of the irrelevant speech items to the participants, not simply their willingness to use looser decision criteria when responding. This result is conceptually equivalent to similar data reported by Osaka et al. (2002) concerning the effect of Reading Span on intrusion errors in a reading task, suggesting that resistance to semantic intrusions might be a general property of working memory capacity rather than anything specific to irrelevant speech or to the auditory modality[1].

It appears from these data that the "standard" ISE is unaffected by working memory capacity, despite the results of Conway et al. (2001) in the dichotic listening paradigm. Rather, that a particular measure of processing the unattended signal is modulated by working memory capacity. It may be noteworthy that both Conway et al.'s (2001) dependent measure, reporting hearing one's own name, and Beaman's (2004) dependent measure, mistakenly reporting an irrelevant speech utterance, require sensitivity to the content of the unattended message. In each case, the message itself is intelligible and meaningful. In previous studies, the meaning of the irrelevant speech had no extra effect upon the distracting effects of speech on serial recall (Buchner et al., 1996) or on counting (Buchner et al., 1998) and the sound need not be intelligible speech in order to create an ISE, provided it contains sufficient token changes (Jones and Macken, 1993; Jones et al., 1990). Thus, both Beaman's (2004) free recall intrusions and Conway et al.'s (2001) cocktail party effect may involve the same kind of inhibitory or interference control function needed to reject intelligible, meaningful speech. This could potentially be related to working memory capacity (c.f., Osaka et al., 2002) yet independent from the control function needed to resist the acoustically-based effects of irrelevant sound more generally.

Hemispheric Differences for Two Types of Unattended Message: Behavioural Evidence

The findings of Beaman (2004) can be characterized as a failure to find the effect of OSPAN upon the ISE predicted on the basis of the dichotic listening literature. Our first inference, that the effects observed in the ISE and dichotic listening paradigms involve different mechanisms is based on the contrast between the positive findings of Conway et al. (2001) and Beaman's (2004) null result. This inference can be tested by looking elsewhere in the behavioural literature for other results indicating a difference in kind between processing of irrelevant sound in dichotic listening and the ISE. Hadlington et al. (2004), investigating the ISE, found the opposite pattern of results to those observed using the dichotic listening technique.

Hadlington et al. (2004) hypothesised that, since language is generally considered to be left lateralized – and contralateral connections for audition are stronger than ipsilateral – irrelevant speech played to the right ear should prove harder to ignore, and hence more disruptive to the primary task, than the same speech played to the left ear. This hypothesis is consistent with data from the dichotic listening paradigm where a right ear advantage for responding to targets is well-documented. For example, there is a right-ear advantage in spotting target-stimuli presented to the right ear in a dichotic listening task (Broadbent and Gregory, 1964; Kimura, 1961, 1967; Studdert-Kennedy and Shankweiler, 1970; see Hugdahl, 2003 for a review of auditory lateralization and dichotic listening studies). The basic idea is that the right-ear advantage in dichotic listening reflects stronger auditory input to the contralateral hemisphere, with the left hemisphere (amongst right-handers) specialized for language processing. This by itself would be sufficient to imply greater processing of input spoken exclusively to the right ear. Additionally, however, the relative strength of the contralateral processing route is enhanced because ipsilateral connections are held to be suppressed, or blocked, by contralateral information. Linguistic information presented to the right hemisphere via the left ear is transferred to the left hemisphere in order to be fully processed, but this transfer is hindered by the simultaneous presence of linguistic information presented to the right ear.

A right-ear disadvantage in the ISE, i.e., a *negative* effect of presenting the irrelevant sound to the right ear, would also be consistent with data from imaging studies indicating a primarily left-hemisphere locus for verbal short-term-memory (e.g., Awh et al., 1996; Henson et al., 2000; Jonides et al., 1998; Logie et al., 2003; Paulesu et al., 1993; Schumacher et al., 1996; Smith et al., 1996, 1998; see Baddeley, 2003 and Henson, 2001 for reviews) as well as linguistic processes generally. For example, Paulesu et al. (1993) and Henson (2001) map verbal components of Baddeley's working memory model onto the left inferior parietal cortex, the left inferior premotor cortex (Broca's region), the left premotor cortex and the right cerebellum regions (although other authors have also made a case for the involvement of further speech-processing regions, see Hickock

[1]Thanks to Steve Lewandowsky for pointing this out.

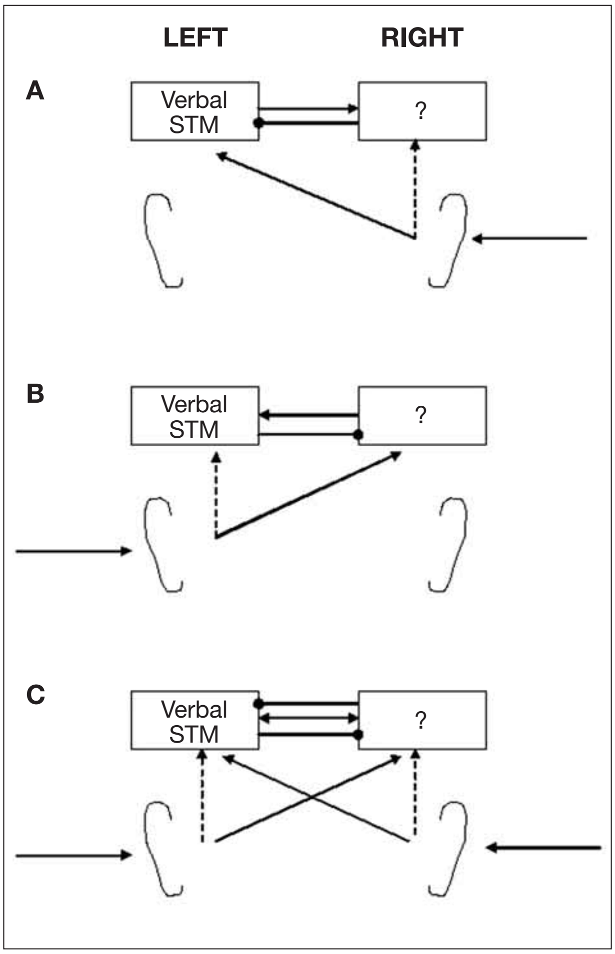

Fig. 1 – *Outline processing system, relative to the supposed primary locus of verbal short-term memory for speech input to left, right and both ears. Arrow heads represent information flow, filled circles represent inhibitory connections. Dotted lines are weak connections and solid lines are strong connections.*

and Buchsbaum, 2001). Irrelevant speech spoken to the right ear would, according to this account, gain obligatory access to the neural locus of verbal short-term memory. Contrastingly, the access route for irrelevant speech spoken to the left ear would be less straightforward because of the relatively weak ipsilateral connections between the left ear and the left hemisphere (Connolly, 1985; Majkowski et al., 1971; Maximilian, 1982; Rosenzweig, 1951). The situation is illustrated in Figure 1, where A shows the access-route of irrelevant speech presented to the right ear only, B shows the access route of irrelevant speech presented to the left ear only and C shows irrelevant speech presented to both ears.

All these data suggest that presentation of irrelevant sound only to the right-ear will create more disturbance to a primary task than presenting irrelevant sound only to the left-ear[2]. Contrary to their initial right ear disadvantage hypothesis, however (greater disruption when irrelevant sound only appears on the right than when it appears only on the left), Hadlington et al. (2004, 2006) observed that speech (the letters B, I, J, N, Z) and non-speech sounds (tones of various pitches) played to the left ear was the more disruptive, as shown in Figure 2. This figure demonstrates that, contrary to the prediction derived from the dichotic listening literature, and from the extant imaging literature, irrelevant speech and non-speech played to the left ear were maximally disruptive to performance on the primary task. The data presented in Figure 2 are taken from the effects of irrelevant sound upon a mental arithmetic task but this pattern of data was replicated using several different dependent variables, including serial recall

[2]Predicting the disturbance provided by the presentation of irrelevant sound to both ears is a little more problematic. The somewhat simplistic neuroanatomical model of auditory lateralization outlined here says little regarding how the experimental situation of messages sent to each ear individually compares to the more realistic situation of the same message sent "free field", i.e., to both ears simultaneously.

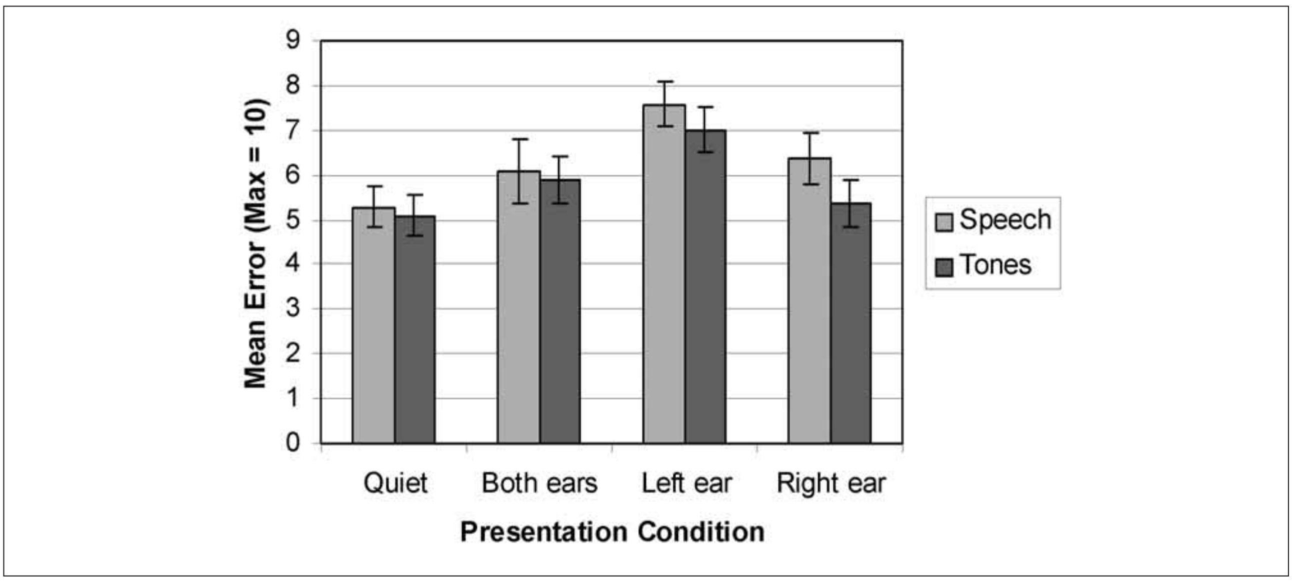

Fig. 2 – *The effects of presenting irrelevant sound (speech or tones) to the left ear, right ear or both ears of participants carrying out a serial recall task. Data taken from Hadlington et al. (2004). Error bars are standard error.*

(Hadlington et al. 2004, 2006). In all cases, presentation to the left ear only was more disruptive than presentation to both ears, suggesting that the activity of the right hemisphere is moderated when both hemispheres receive the same message.

These data support our inference that the processing of unattended speech undertaken in dichotic listening tasks differs from that apparent in the irrelevant sound paradigm. The initial contention that ISE and dichotic listening investigate the same aspects of processing unattended speech is thrown into doubt by the failure to find the ISE results predicted on the basis of dichotic listening investigations (Beaman, 2004) and the Hadlington et al. (2004) finding of a diametrically opposite pattern of results in ISE (extensive processing of left ear input) from that seen in dichotic listening (the right ear advantage). In an attempt to reconcile findings from the two paradigms we turn now to a second source of information, neuroimaging studies of passive listening and, in particular positron emission tomography (PET) studies of listening in noise.

Hemispheric Differences for Two Types of Unattended Message: Imaging Data

Two previous neuroimaging studies examined the neural locus of the ISE using PET (Gisselgård et al., 2003, 2004). In Gisselgård et al.'s (2003) study visually presented to-be-recalled stimuli and to-be-ignored irrelevant speech were presented simultaneously. The baseline for comparison with serial recall was a digit repetition task. Robust activation appeared bilaterally in the superior temporal region when speech conditions were compared to quiet. Comparisons of repeated speech ('steady-state') with varying speech ('changing-state') were also made because, as previously noted, an ISE occurs whenever abrupt changes in pitch or other physical characteristics (Jones et al., 1992, 1993) are present within the acoustic signal. With regard to speech stimuli, this means that varied utterances (e.g., 'A-B-C') reliably produce an ISE but repeated utterances (e.g., 'A-A-A') do not. Pitch-shifted tones also reliably show an effect, whereas steady-state (same pitch) sequences do not (Jones and Macken, 1993). This phenomenon, known as the "changing-state effect", has been considered the signature effect of irrelevant sound (Jones et al., 2004).

Changing *versus* steady-state speech comparisons in Gisselgård et al.'s (2003) study showed a significant decrease in activation in the left superior temporal cortex and a weaker but still significant decrease in left inferior parietal cortex, bilateral secondary auditory and inferior/middle frontal areas. These decreases were associated with changing-state irrelevant speech in the context of a serial recall task compared to steady-state irrelevant speech in the same context. Gisselgård et al. (2003) note that decrease in activity is greater in the left than the right superior temporal region and suggest that the relatively weak activation of the left inferior parietal cortex (Brodmann's area 40) in serial recall relative to digit repetition may be related to a general inhibitory influence of changing-state irrelevant speech (pp. 1906-1908). These results are interesting but difficult to interpret as relative decreases in activation, rather than the more usual increases, are reported. The task of interpretation is made no easier because in only one of Gisselgård et al.'s (2004) studies was an ISE observed in the behavioural data and, in this study, additional activation was observed in the prefrontal cortex. This additional activity may, however, have been a consequence of the more difficult primary memory task participants were asked to perform (Campbell, 2005).

In addition to these studies by Gisselgård et al. (2003, 2004), PET-studies on speech perception in the presence of irrelevant, competing speech are informative. The comparative simplicity of such studies makes the neural activity associated with processing irrelevant speech easier to interpret and, arguably, more informative than the more complex situations examined by Gisselgård et al. (2003, 2004). Hickock and Buchsbaum (2001) argue that brain regions involved in speech perception, and in particular the temporal lobes, should be considered part of the verbal working memory network in addition to other areas already implicated (e.g., Paulesu et al., 1993) and there is no reason to suppose the same systems are not also employed for processing of unattended speech. When these speech perception systems are examined, Hadlington et al.'s (2004) data become more comprehensible.

In studies of the neural basis of speech perception, the left superior temporal lobe has been found to respond to intelligible speech (Buchsbaum et al., 2001; Narain et al., 2003; Scott et al., 2000) and to discrimination of phonetic structure (Zatorre et al., 1992). This is consistent with Hadlington et al.'s (2004) original hypothesis that presentation of irrelevant sound to the right ear would give the greatest ISE (see Figure 1). In addition, however, the right superior temporal lobe has been found to respond to signals with dynamic pitch variation, regardless of intelligibility (Patterson et al., 2002; Scott et al., 2000; Zatorre et al., 1992). If the same applies to unattended speech, this creates a framework in which the data presented by Beaman (2004) and Hadlington et al. (2004) can be brought together and interpreted. Changing-state irrelevant speech gives rise to an ISE regardless of the intelligibility of the speech (Jones et al., 1996). The notion of changing-state, however, implies significant dynamic pitch variation which steady-state sound by definition does not possess. Dynamic pitch variation in normal speech perception is associated with the right superior temporal gyrus, and the right hemisphere receives its strongest input from the left ear, which also happens to produce the largest ISE. However, these speculations are based imaging data obtained from normal speech perception, not from neural processing of to-be-ignored speech.

Recent studies by Scott et al. (2004 submitted) have investigated the neural activity associated with unattended speech. The first experiment examined cerebral activity when participants were asked to follow a female speaker in the presence of either continuous noise or a further (male) speaker (Scott et al., 2004). Listening to the to-be-attended speech in the presence of to-be-ignored speech resulted in extensive bilateral activation relative to listening to the to-be-attended speech in the presence of to-be-ignored continuous noise. Bilateral activation is consistent with similar results obtained by Gisselgård et al. (2003) for their comparison of changing- and steady-state speech with quiet, although Gisselgård et al.'s (2003, 2004)result is contaminated by the presence in the data of activity associated with the digit serial recall task. Unfortunately, in Scott et al.'s (2004) study the bilateral activation is also open to multiple interpretations. It could be interpreted in terms of semantic analysis of the to-be-ignored speech, acoustic analysis of the same, or a confound wherein silent "gaps" in the to-be-ignored signal afford clear glimpses of the to-be-attended stimuli (Festen and Plomp, 1990). No such glimpses were possible in the continuous noise control condition. Since the results of functional imaging studies (PET and fMRI) are entirely dependent on the choice of baseline, a further study was conducted to investigate the influence of lexico-semantic *versus* acoustic processing of the masking speech, and to control for the possibility of low level glimpses of the target signal.

Scott et al. (submitted) conducted a study in which the to-be-attended (female) speaker appeared concurrently with either a to-be-ignored male speaker, to-be-ignored spectrally rotated male speech (Blesser, 1972) or to-be-ignored signal correlated noise versions of the male speaker. In this study, all to-be-ignored stimuli were amplitude-modulated, thereby allowing glimpses of the target to-be-attended speech in all conditions. Rotated speech is used as a baseline stimulus for speech perception studies, since it has the same spatio-temporal structure and pitch variation as unrotated speech (Narain et al., 2003; Scott et al., 2000). Spectrally-rotated speech thus preserves the acoustic complexity and pitch of the original speech, but not the meaning. Crucially, spectrally rotated speech therefore preserves the acoustic features of pitch variation which pertain to "changing states" in the signal, without a confound of lexical activations. Contrasting spectrally-rotated speech with signal-correlated noise enables the

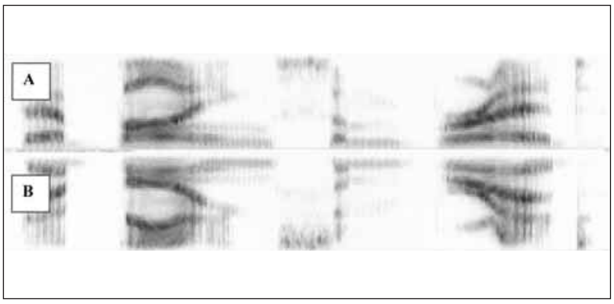

Fig. 3 – *An example of spectrally-rotated speech. The upper panel (A) shows forward, intelligible speech, the lower panel (B) shows the same speech when spectrally rotated. Further examples can be viewed and heard at: http://www.phon.ucl.ac.uk/home/brain/.*

distinction between neural responses for to-be-ignored pitch properties of speech, and those for the to-be-ignored lexico-semantic content of speech.

A comparison of normal speech and spectrally-rotated speech is shown in Figure 3 (see Scott et al., 2000, for technical details of spectral rotation). As expected, there was bilateral activation of the superior temporal gyrus (STG) associated with listening to speech-in-speech minus speech-in-signal-correlated noise (see Figure 4, panel A),

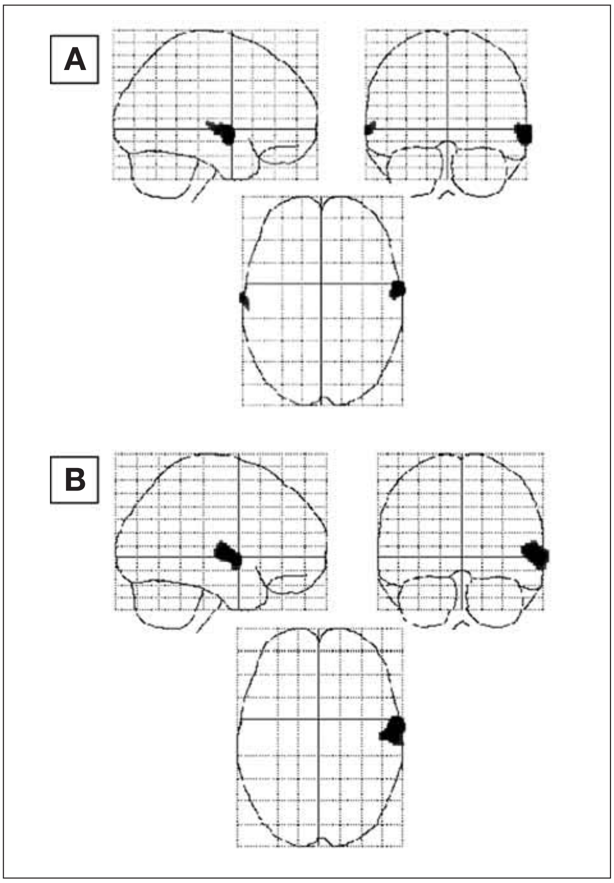

Fig. 4 – *A glass brain representation of the neural activity associated with masking speech. The upper panel (A) shows the bilateral activation observable with intelligible, forward speech. The lower panel (B) shows the right hemisphere activation observable when the speech is spectrally rotated and no longer either intelligible or meaningful. Data taken from Scott et al. (submitted).*

indicating that the message spoken by the unattended male speaker was receiving considerably more processing than unattended signal-correlated noise. Thus, a potential artefact in Scott et al.'s (2004) data can be ruled out as any quiet periods when only the attended voice was present occurred equivalently in the two experimental and the (signal-correlated) control condition.

The most interesting condition is that examining spectrally-rotated speech minus signal-correlated noise. In this condition, it is possible to isolate the effects of an unattended speech signal, matched with speech in pitch variation, but devoid of intelligibility or semantic content. Any neural activity in this condition must therefore be associated with processing the pitch dynamics of the unintelligible to-be-ignored speech. When this analysis is carried out, a peak of activation in the right STG is observed (Figure 4, panel B). In other words, consistent with the left ear disadvantage reported by Hadlington et al., there is considerable processing of the irrelevant signal occurring in the right hemisphere. These data are consistent not only because both suggest a role for the right hemisphere in processing to-be-ignored speech but, importantly, the speech has been rendered unintelligible by spectral rotation. This processing is therefore tied to the pitch change, rather than the lexico-semantics, of the unattended speech. The inference we draw from this is that the right hemisphere processes acoustic properties contributing to the disruptive, "changing-state" nature of irrelevant sound, producing the left-ear disadvantage observed by Hadlington et al. (2004). The stimuli used by Hadlington et al. (2004) were changing-state speech (the letters B, I, J, N, Z spoken in a male voice) and changing-state non-speech (tones varying in pitch) and therefore, in common with previous observations, created an ISE despite containing little or no semantic content (Buchner et al., 1996; Jones and Macken, 1993). Steady-state irrelevant sound (a single repeated utterance or tone) produced no ISE and no right-ear disadvantage (Hadlington et al., 2006).

Parallel Routes for Processing Unattended Speech in Dichotic Listening and the Ise

The conclusion from the imaging data is that the right hemisphere is sensitive to pitch variation in unattended speech, but plays no role in analysing the meaning of unattended speech. In conjunction with the other studies already discussed, this conclusion can be taken further. As the right hemisphere processes the pitch variation in speech, this must be the source of the enhanced ISE observed by Hadlington et al. (2004). These data fit well with existing results concerning the ISE as it is already established that non-speech sound creates an ISE (Jones and Macken, 1993) but only usually when the sound is subject to abrupt changes in pitch or other physical characteristics (Jones et al., 1992, 1993). The acoustic characteristics of spectrally-rotated speech processed by the right hemisphere fit the requirements for "changing-state" information. However, there is no necessary relationship between the analysis of changing-state information identified as taking place in the right hemisphere and working memory capacity, as established by Beaman (2004), Elliott and Cowan (2005) and Neath and Surprenant (2004). Either activity in right STG is "informationally encapsulated" (Fodor, 1983) and cannot be moderated by the attentional mechanisms underlying working memory span, or the output of the right STG is similarly resistant to top-down attentional influences of the kind represented by working memory span. This leaves the relationship between working memory span and dichotic listening tasks to be explained.

Activation of the right hemisphere is not associated with lexico-semantic processing in Scott et al.'s (2004, submitted) study. Therefore, the appearance of the "standard" ISE, for which unintelligible changing-state acoustic information is sufficient, is neuroanatomically as well as behaviourally distinct from the lexico-semantic effects of irrelevant speech. Semantic effects of irrelevant speech are not necessary to produce an ISE (Jones et al., 1996). Since the left hemisphere responds to the meaning of speech, the final inference is that this is most probably the source of the right-ear advantage in dichotic listening, which normally takes the form of an advantage for spotting or recalling intelligible and relatively meaningful targets (single words or sequences of words, respectively). Conway et al.'s (2001) results with dichotic listening and Beaman's (2004) results with meaningful irrelevant speech reviewed earlier (Table I) suggest that activity within this system may also be moderated by working memory span, unlike the acoustic analysis taking place within the right hemisphere. Thus, it seems likely that the either the activity or the output of the lexico-semantic analysis system identified with the left STG is not informationally encapsulated and it is open to moderation by attentional factors reflected in working memory span in a way that dynamic pitch analysis of the irrelevant sound is not. The situation as we have analysed it so far is depicted graphically in Figure 5.

These data now connect a number of experimental paradigms used to investigate similar phenomena that have been pursued more-or-less independently by researchers interested in auditory attention (the dichotic listening paradigm) working memory (the ISE) and speech perception in noisy environments. The data cross two separate regions of investigation (dichotic listening and the ISE) and

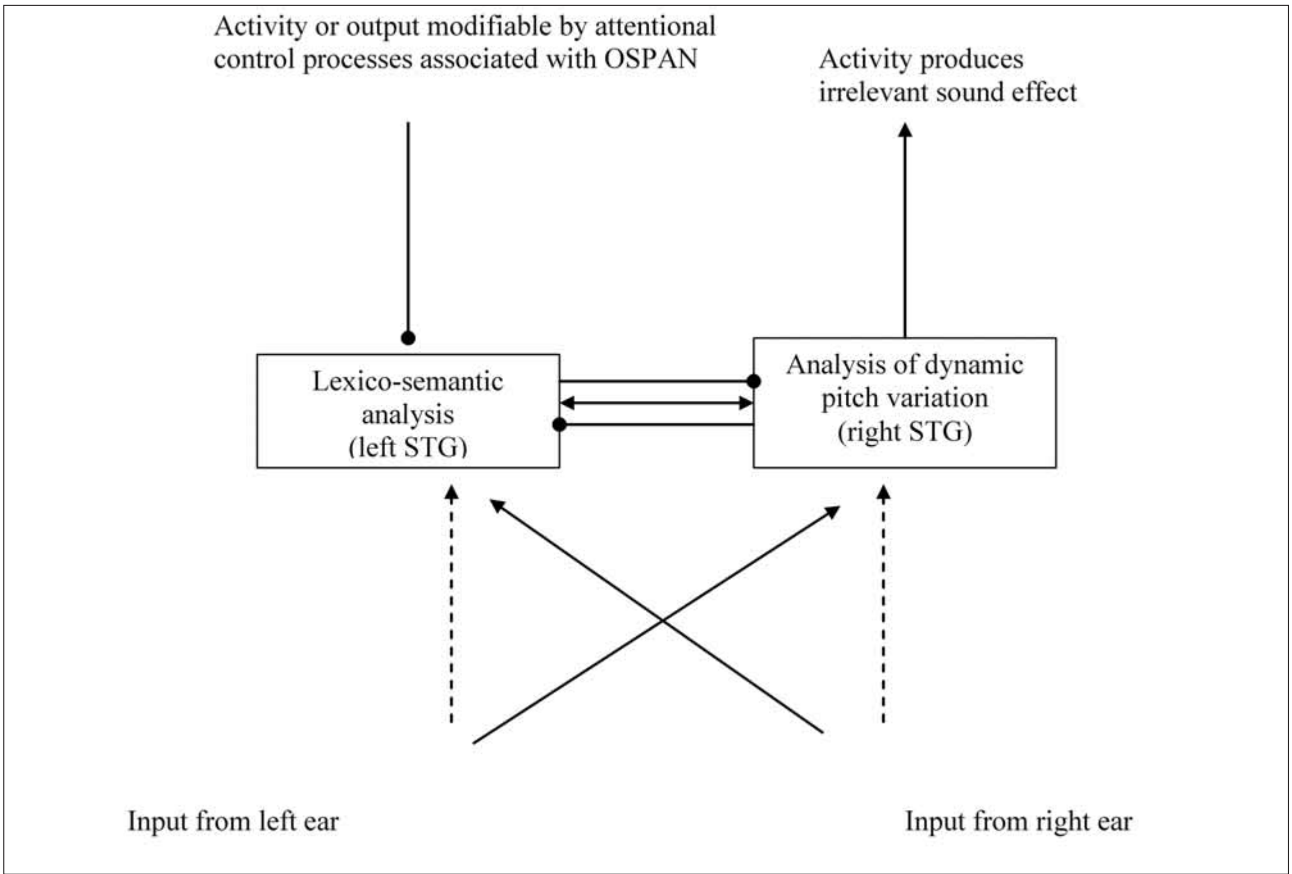

Fig. 5 – *A schematic showing the proposed analysis routes for irrelevant, unattended sound. Arrows indicate interactivity of processing. Arrow heads represent information flow, filled circles represent inhibitory connections. Dotted lines are weak connections and solid lines are strong connections. This schematic represents only the processing routes for which we have direct evidence. Other possible routes (e.g., acoustic processing occurring in the left hemisphere) are not depicted.*

two separate experimental tools (experimental psychology and functional imaging). We conclude that unattended speech is processed along at least two neural pathways, resulting in entirely different behavioural effects, a finding that has not previously been acknowledged. Although we maintain that this discovery contributes a great deal to our understanding of the analysis of unattended or irrelevant auditory messages there are a number of important issues that remain unresolved. From the point of view of the ISE, the most salient unresolved issue is where in the system the irrelevant sound interference actually occurs. A plausible hypothesis that the interference occurs between two sets of lexical analyses (the to-be-recalled and to-be-ignored stimuli) co-occurring in the same location (i.e., the left hemisphere) seems unlikely. However the current data stop short of indicating exactly where the ISE interference occurs (in either cognitive or neural terms). It is plausible that multiple effects occur and there is no single locus of interference, although the available data implicate activity within the right STG as a necessary part of the process. The Beaman (2004) and Conway et al. (2001) results suggest, however, that lexico-semantic effects of irrelevant speech are distinguishable from solely acoustic effects in terms of their cognitive profile and also, as we have indicated, in terms of their neural origin. Lexico-semantic effects are related to OSPAN (Table I) where acoustic-only interference effects are not.

References

AWH E, JONIDES J, SMITH EE, SCHUMACHER EH, KOEPPE RA and KATZ S. Dissociation of storage and rehearsal in verbal working memory. *Psychological Science, 7:* 25-31, 1996.

BADDELEY AD. *Working Memory*. Oxford: Clarendon Press, 1986.

BADDELEY AD. Working memory: Looking back and looking forwards. *Nature Reviews Neuroscience, 4:* 829-839, 2003.

BADDELEY AD and WEISKRANTZ L. *Attention: Selection, Awareness and Control. A tribute to Donald Broadbent.* Oxford: Clarendon Press, 1993.

BEAMAN CP. The irrelevant sound phenomenon revisited: What role for working memory capacity? *Journal of Experimental Psychology: Learning, Memory and Cognition, 30:* 1106-1118, 2004.

BEAMAN CP and JONES DM. The role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23:* 459-471, 1997.

BEAMAN CP and JONES DM. Irrelevant sound disrupts order information in free recall as in serial recall. *Quarterly Journal of Experimental Psychology, 51A:* 615-636, 1998.

BLESSER B. Speech perception under conditions of spectral transformation: I. Phonetic characteristics. *Journal of Speech and Hearing Research, 15:* 5-41, 1972.

BRIDGES AM and JONES DM. Word-dose in the disruption of serial recall by irrelevant speech: Phonological confusions or changing state? *Quarterly Journal of Experimental Psychology, 49A:* 919-939, 1996.

BROADBENT DE. *Perception and Communication*. Oxford: Pergamon Press, 1958.

BROADBENT DE and GREGORY M. Accuracy and recognition for speech presented to the right and left ears. *Quarterly Journal of Experimental Psychology, 16:* 359-360, 1964.

BUCHNER A and ERDFELDER E. Word frequency of irrelevant speech distractors affects serial recall. *Memory and Cognition, 33:* 86-97, 2005.

BUCHNER A, IRMEN L and ERDFELDER E. On the irrelevance of semantic information for the "irrelevant speech" effect. *Quarterly Journal of Experimental Psychology, 49A:* 765-779, 1996.

BUCHNER A, ROTHERMUND, K WENTURA D and MEHL B. Valence of distractor words increases the effects of irrelevant speech on serial recall. *Memory and Cognition, 32:* 722-731, 2004.

BUCHNER A, STEFFENS MC, IRMEN L and WENDER KF. Irrelevant auditory material affects counting. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24:* 48-67, 1998.

BUCHSBAUM BR, HICKOK G and HUMPHRIES C. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science, 25:* 663-678, 2001.

CAMBELL T. The cognitive neuroscience of auditory distraction. *Trends in Cognitive Sciences, 9:* 3-5, 2005.

CHERRY EC. Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America, 25:* 975-979, 1953.

CONNOLLY JF. Stability of pathway-hemispheric differences in the auditory event related potential (ERP) to monaural stimulation. *Psychophysiology, 22:* 87-96, 1985.

CONWAY AR, COWAN N and BUNTING MF. The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review, 8:* 331-335, 2001.

CONWAY AR and ENGLE RW. Working memory and retrieval: A resource dependent inhibition model. *Journal of Experimental Psychology: General, 123:* 354-373, 1994.

COWAN N. *Attention and Memory: An Integrated Framework*. Oxford: Clarendon Press, 1995.

COWAN N and BARRON A. Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception and Psychophysics, 41:* 393-401, 1987.

DANEMAN M and CARPENTER PA. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19:* 450-466, 1980.

DRIVER J and BAYLIS GC. Cross-modal negative priming and interference in selective attention. *Bulletin of the Psychonomic Society, 31:* 45-48, 1993.

ELLERMEIER W and ZIMMER K. Individual differences in susceptibility to the "irrelevant speech effect". *Journal of the Acoustical Society of America, 102:* 2191-2199, 1997.

ELLIOTT EM and COWAN N. Coherence of the irrelevant sound effect: Individual profiles of short-term memory and susceptibility to task-irrelevant materials. *Memory and Cognition, 33:* 761-767, 2005.

ELLIOTT EM, COWAN N and VALLE-INCLAN F. The nature of cross-modal color-word interference effects. *Perception and Psychophysics, 60:* 761-767, 1998.

ENGLE RW. What is working memory capacity? In Roediger HL III, Nairne JS, Neath I and Surprenant AM (Eds), *The Nature of Remembering: Essays in Honor of Robert G. Crowder*. Washington D.C.: APA, 2001.

ENGLE RW, CONWAY AR, TUHOLSKI SW and SHISLER RJ. A resource account of inhibition. *Psychological Science, 6:* 122-125, 1995.

ENGLE RW, KANE MJ and TUHOLSKI SW. Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In Miyake A and Shah P (Eds), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.* Cambridge: Cambridge University Press, 1999.

FESTEN J and PLOMP R. Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America, 88:* 1725-36, 1990.

FODOR J. *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

GISSELGÅRD J, PETERSSON KM, BADDELEY A and INGVAR M. The irrelevant speech effect: A PET study. *Neuropsychologia, 41:* 1899-1911, 2003.

GISSELGÅRD J, PETERSSON KM and INGVAR M. The irrelevant speech effect and working memory load. *NeuroImage, 22:* 1107-1116, 2004.

HADLINGTON L, BRIDGES AM and BEAMAN CP. A left-ear disadvantage for the presentation of irrelevant sound: Manipulations of task requirements and changing state. *Brain and Cognition, 61:* 159-171, 2006.

HADLINGTON L, BRIDGES AM and DARBY RJ. Auditory location in the irrelevant sound effect: The effects of presenting auditory stimuli to either the left ear, right ear or both ears. *Brain and Cognition, 55:* 545-557, 2004.

HENSON RN. Neural working memory. In Andrade J (Ed), *Working Memory in Perspective.* Hove: Psychology Press, 2001.

HENSON RN, BURGESS N and FRITH CD. Recoding, storage, rehearsal and grouping in verbal short-term memory: An fMRI study. *Neuropsychologia, 38:* 426-440, 2000.

HICKOK G and BUCHSBAUM B. Temporal lobe speech perception systems are part of the verbal working memory circuit: Evidence from two recent fMRI studies. *Behavioral and Brain Sciences, 26:* 740-741, 2001.

HUGDAHL K. Dichotic listening in the study of auditory laterality. In Hugdhal K and Davidson RJ (Eds), *The Asymmetrical Brain*. Cambridge, MA: MIT Press, 2003.

HUGHES RW and JONES DM. A negative order repetition priming effect: Inhibition of order in unattended sequences? *Journal of Experimental Psychology: Human Perception and Performance, 29:* 199-218, 2003.

HUGHES RW and JONES DM. The impact of order incongruence between a task-irrelevant auditory sequence and a task-relevant visual sequence. *Journal of Experimental Psychology: Human Perception and Performance, 31:* 316-327, 2005.

JONES DM, BEAMAN CP and MACKEN WJ. The object-oriented episodic record model. In Gathercole SE (Ed), *Models of Short-Term Memory*. Hove: Psychology Press, 1996.

JONES DM and MACKEN WJ. Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19:* 369-381, 1993.

JONES DM and MACKEN WJ. Organizational factors in the effect of irrelevant speech: The role of spatial location and timing. *Memory and Cognition, 23:* 192-200, 1995.

JONES DM, MACKEN WJ and MURRAY AC. Disruption of visual short-term memory by changing-state auditory stimuli: The role of segmentation. *Memory and Cognition, 21:* 318-328, 1993.

JONES DM, MACKEN WJ and NICHOLLS AP. The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory and Cognition, 30:* 656-674, 2004.

JONES DM, MADDEN C and MILES C. Privileged access by irrelevant speech to short-term memory: The role of changing-state. *Quarterly Journal of Experimental Psychology, 44A:* 645-669, 1992.

JONES DM, MILES C and PAGE J. Disruption of proof-reading by irrelevant speech: Effects of attention, arousal or memory? *Applied Cognitive Psychology, 4:* 89-108, 1990.

JONIDES J, SCHUMACHER EH, SMITH EE, KOEPPE RA, AWH E, REUTER-LORENZ PA, MARSHUETZ C and WILLIS CR. The role of parietal cortex in verbal working memory. *The Journal of Neuroscience, 18:* 5026-5034, 1998.

KANE MJ and ENGLE RW. Working memory capacity, proactive interference and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26:* 336-358, 2000.

KIMURA D. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology, 15:* 166-171, 1961.

KIMURA D. Functional asymmetry of the brain in dichotic listening. *Cortex, 3:* 163-168, 1967.

LECOMPTE DC. Extending the irrelevant speech effect beyond serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20:* 1396-1408, 1994.

LOGIE RH, VENNERI A, DELLA SALA S, REDPATH TW and MARSHALL I. Brain activation and the phonological loop: The impact of rehearsal. *Brain and Cognition, 53:* 293-296, 2003.

MACKEN WJ, TREMBLAY S, HOUGHTON RJ, NICHOLLS AP and JONES DM. Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 29:* 43-51, 2003.

MAJKOWSKI J, BOCHENEK Z, BOCHENEK W, KNAPIK-FIJALKOWSKA D and KOPEC J. Latency of average evoked potentials to contralateral and ipsilateral auditory stimulation in normal subjects. *Brain Research, 24:* 416-419, 1971.

MAXIMILIAN VA. Cortical blood flow asymmetry during monaural verbal stimulation. *Brain and Language, 15:* 1-11, 1982.

MORAY N. Attention in dichotic listening: Affective cues and the

influence of instructions. *Quarterly Journal of Experimental Psychology, 11:* 56-60, 1959.
NARAIN C, SCOTT SK, WISE RJ, ROSEN S, LEFF A, IVERSEN SD and MATTHEWS PM. Defining a left-lateralized response specific to intelligible speech using fMRI. *Cerebral Cortex, 13:* 1362-1368, 2003.
NEATH I, FARLEY LA and SURPRENANT A. Directly assessing the relationship between irrelevant speech and articulatory suppression. *Quarterly Journal of Experimental Psychology, 56A:* 1269-1278, 2003.
NEATH I and SURPRENANT A. *Individual Differences and the Irrelevant Speech Effect*. Unpublished manuscript, 2004.
NEELY CB and LECOMPTE DC. The importance of semantic similarity to the irrelevant speech effect. *Memory and Cognition, 27:* 37-44, 1999.
NICHOLLS AP and JONES DM. Capturing the suffix: Cognitive streaming in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28:* 12-28, 2002.
OSAKA M, NISHIZAKI Y, KOMORI M and OSAKA N. Effect of focus on verbal working memory: Critical role of the focus word in reading. *Memory and Cognition, 20:* 562-571, 2002.
OSWALD CJ, TREMBLAY S and JONES DM. Disruption of comprehension by the meaning of irrelevant speech. *Memory, 8:* 345-350, 2000.
PATTERSON RD, UPPENKAMP S, JOHNSRUDE IS and GRIFFITHS TD. The processing of temporal pitch and melody information in auditory cortex. *Neuron, 36:* 767-776, 2002.
PAULESU E, FRITH CD and FRACKOWIAK RS. The neural correlates of the verbal component of working memory. *Nature, 362:* 342-344, 1993.
ROSEN VM and ENGLE RW. Working memory capacity and suppression. *Journal of Memory and Language, 39:* 418-436, 1989.
ROSENZWEIG MR. Representations of the tow ears at the auditory cortex. *American Journal of Physiology, 167:* 147-214, 1951.
SCHUMACHER EH, LAUBER E, AWH E, JONIDES J, SMITH EE and KOEPPE RA. PET evidence for an amodal verbal working memory system. *NeuroImage, 3:* 79-88, 1996.
SCOTT SK, BLANK CC, ROSEN S and WISE RJ. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123:* 2400-2406, 2000.
SCOTT SK, ROSEN S, DAVIS J, BEAMAN CP and WISE RJ. The neural processing of unattended speech: dual mechanisms in the cocktail party effect? Submitted.
SCOTT SK, ROSEN S, WICKHAM L and WISE RJ. A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. *Journal of the Acoustical Society of America, 115:* 813-821, 2004.
SMITH EE, JONIDES J and KOEPPE RA. Dissociating verbal and spatial working memory using PET. *Cerebral Cortex, 6:* 11-20, 1996.
SMITH EE, JONIDES J, MARSHUETZ C and KOEPPE RA. Components of verbal working memory: Evidence from neuroimaging. *Proceedings of the National Academy of Sciences of the USA, 95:* 876-882, 1998.
STUDDERT-KENNEDY M and SHANKWEILER DP. Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America, 48:* 579-594, 1970.
TURNER ML and ENGLE RW. Is working memory capacity task-dependent. *Journal of Memory and Language, 28:* 127:154, 1989.
WOOD N and COWAN N. The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory and Cognition, 2:* 255-260, 1995.
ZATORRE RJ, EVANS AC, MEYER E and GJEDDE A. Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256:* 846-849, 1992.

Philip Beaman, Department of Psychology, School of Psychology and Clinical Language Sciences, University of Reading, Earley Gate, Whiteknights, Reading, RG6 6AL, UK. e-mail: c.p.beaman@reading.ac.uk

(*Received 1 March 2005; reviewed 6 June 2005; revised 12 July 2005; accepted 21 October 2005*)