Available online at www.sciencedirect.com
ScienceDirect
Clinical Psychology Review 27 (2007) 140–154
CLINICAL PSYCHOLOGY REVIEW

Neural correlates of inner speech and auditory verbal hallucinations: A critical review and theoretical integration

Simon R. Jones *, Charles Fernyhough

Department of Psychology, Durham University, South Road, Durham, DH1 3LE, UK

Received 11 May 2006; received in revised form 15 August 2006; accepted 2 October 2006

Abstract

The neuroimaging and neurophysiological literature on inner speech in healthy participants and those who experience auditory verbal hallucinations (AVHs) is reviewed. AVH-hearers in remission and controls do not differ neurologically on tasks involving low levels of verbal self-monitoring (VSM), such as reciting sentences in inner speech. In contrast, on tasks involving high levels of VSM, such as auditory verbal imagery, AVH-hearers in remission show less activation in areas including the middle and superior temporal gyri. This pattern of findings leads to a conundrum, given that mentation involving low levels of VSM is typically held to form the raw material for AVHs. We address this by noting that existing neuroimaging and neurophysiological studies have been based on unexamined assumptions about the form and developmental significance of inner speech. We set out a Vygotskian approach to AVHs which can account for why they are generally experienced as the voice of another person, with specific acoustic properties, and a tendency to take the form of commands. On this approach, which we argue is consistent with the neural correlates evidence, AVHs result from abnormalities in the transition between condensed and expanded dialogic inner speech. Further potential empirical tests of this model are discussed.

Keywords: Auditory verbal hallucinations; Dialogue; Inner speech; Verbal self-monitoring; Vygotsky

Contents

* Corresponding author. Tel.: +44 1913 343240; fax: +44 1913 343241.
E-mail address: s.r.jones@durham.ac.uk (S.R. Jones).

0272-7358/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cpr.2006.10.001

Causes of auditory verbal hallucinations (AVHs), the experience of perceiving speech without corresponding external stimulation (Stephane, Barton, & Boutros, 2001), are still not well understood. Any theory claiming to offer an explanation of AVHs must be able to explain why a voice is heard in the absence of an external stimulus. In addition, any such account must explain why such a voice is experienced as generated or authored by an individual other than one's own self. A complete theory of AVHs must further be able to account for key features of the phenomenology of the experience. While the phenomenology of AVHs varies widely between different voice-hearers, broad themes can be extracted. One issue concerns the personal identity of the voice heard. While 28% of voice-hearers in a sample interviewed by Leudar, Thomas, McNally, and Glinski (1997) reported hearing only incognito voices, most (64%) reported hearing a voice which they could clearly identify as being that of a family member or acquaintance. Others, notably only those with a diagnosis of schizophrenia, reported hearing voices belonging to public figures.

Surveys of the numbers of different voices heard by voice-hearers have found an average of 2–3 voices (Leudar et al., 1997; Nayani & David, 1996). Ninety-six percent of voice-hearers in Leudar et al.'s (1997) sample reported being the target of the voice, which spoke to them rather than addressing another voice or another person. This highlights the fact that issues of the content and pragmatics of voices must also be addressed by a complete theory. Voices can attempt to perform a wide range of functions, including advising on possible actions, requesting specific actions, and inhibiting actions. However, AVHs most commonly take the form of commands, for example, 'Get the milk' or 'Go to the hospital' (Nayani & David, 1996). These command AVHs were reported by 84% of voice-hearers in Nayani and David's phenomenological survey and appear common in the autobiographical literature too. For example, North (1990) recalls how, during the period of her life when she was diagnosed with schizophrenia, she heard distinct voices issuing commands such as, 'Be good', 'Do bad', 'Stand up', and 'Sit down' (p. 60).

In summary, any satisfactory account of AVHs must be able to explain the following aspects of the experience: (1) why the AVH-hearer experiences a voice in the absence of any external stimulus; (2) why this voice is experienced as generated/authored by an individual other than one's own self; (3) why the voice is often perceived as having person-specific (e.g., acoustic) properties, distinct from one's own; and (4) why AVHs have characteristic contents and pragmatics, particularly their tendency to appear as second-person assertive utterances (e.g., commands). To date, no theory has been entirely successful in accounting for all of these aspects of the phenomenon, with the third and fourth criteria, in particular, rarely addressed. Our aim in this article is to present a theoretical framework which, we suggest, is sufficiently broad to do justice to the rich and varied phenomenology of AVHs. In doing so, we attempt what Bentall (1990) recommends a scientific theory of hallucinations should do, namely, "explain the experiences of the hallucinator in terms of underlying cognitive mechanisms" (p. 83).

In setting out our arguments, we will be drawing upon an assumption which has guided much thinking on AVHs, namely that they are a form of inner speech (Bentall, 2003; Bick & Kinsbourne, 1987; Fernyhough, 2004; Jones & Fernyhough, in press; Leudar & Thomas, 2000; Maudsley, 1886; Seal, Aleman, & McGuire, 2004). A *prima facie* case for the involvement of inner speech in AVHs concerns a basic commonality between the two kinds of experience, namely that both involve some form of internal verbal mentation, or 'voice in the head.' Furthermore, like inner speech, AVHs are often relevant to the voice-hearer's ongoing activities, and may attempt to comment on or regulate behavior (Leudar & Thomas, 2000). While other explanations of AVHs are possible (e.g., Hoffman & McGlashan, 1997; Lennox et al., 2000), the strong conceptual and empirical support for a relation between the two phenomena suggests that considering AVHs as a disorder of inner speech remains a fruitful line of enquiry. In this article, we propose that further light may be shone on the puzzle of AVHs through a more careful investigation of the nature and developmental significance of inner speech. Our arguments will be informed by the claims of L. S Vygotsky (1934/1987) that inner speech can best be understood in terms of its being the endpoint of a developmental process. A guiding assumption, which we will argue is supported by empirical evidence from the study of children's speech, is that inner speech in healthy individuals is irreducibly dialogic in nature. We contend that AVHs result from an abnormality in the typical process of inner speech production, specifically the transition between condensed and expanded varieties of inner speech. These ideas require substantial unpacking and will now be more fully explicated.

1. Neuroimaging of inner speech and AVHs

Those working within the cognitive neuropsychology tradition typically view experiences such as AVHs as being underpinned by specific cognitive deficits or biases, which in turn must have neural instantiation. As remarked, many authors have proposed that the cognitive deficit involved in AVHs is a disorder of inner speech. Such a proposal is

supported by evidence that brain areas implicated in inner speech, such as the left inferior frontal gyrus (including Broca's area) and the right temporal cortex (Huang, Carr, & Cao, 2001; Jones, in press; McGuire, Silbersweig, Murray, et al., 1996) are activated during AVHs (McGuire, Shah, & Murray, 1993; Shergill et al., 2004; Shergill, Brammer, Williams, Murray, & McGuire, 2000). Although some have questioned whether the finding of inferior frontal gyrus activation during AVHs is an artifact of using a button-press to signal onset of AVHs during imaging (Hunter & Spence, 2005), other studies which have not used a button-press to signal onset, but have instead asked post-imaging whether an AVH was experienced, have also shown inferior frontal gyrus activation in AVHs (Shergill, Brammer, et al., 2000). This said, several other imaging studies have failed to replicate the finding that inferior frontal gyrus activation is associated with the experience of AVHs (e.g., Lennox et al., 2000; Silbersweig et al., 1995). For the present we will make the working assumption that the neural substrates underlying inner speech are implicated in the genesis of AVHs. We will return to offer an explanation for these seemingly contradictory findings about the involvement of the inferior frontal gyrus later in this article.

Several specific proposals have been put forward within a cognitive neuropsychological framework to understand how atypical processing of inner speech may cause AVHs (Frith, Rees, & Friston, 1998; Jones & Fernyhough, in press; Seal et al., 2004). At the base of many of these proposals is the neuropsychological mechanism originally proposed by Frith and Done (1988), in which faulty monitoring of inner speech results in verbal thoughts not being recognized as one's own and being perceived instead as someone else's. The cognitive capacity responsible for monitoring inner speech is termed verbal self-monitoring (VSM). VSM has been proposed to operate through a corollary discharge system whereby discharges from the motor speech-producing areas in the frontal lobes 'warn' the auditory cortex that self-produced speech is about to be produced by dampening its response (Creutzfeldt, Ojeman, & Lettich, 1989; Ford & Mathalon, 2005; Martikainen, Kaneko, & Hari, 2005). Such a system may signal through the white matter tracts of the arcuate fasciculus which has pathways (among others) that originate in Broca's area and project to Wernicke's area (Hubl et al., 2004).

VSM can also be construed as entailing the appraisal of what one has said in relation to what one intended to say (McGuire, Silbersweig, & Frith, 1996). McGuire et al. (1996) used a PET imaging study to attempt to pin down the brain regions involved in VSM, by experimentally manipulating auditory verbal feedback while participants read aloud, and hence inducing a disparity between what the participant expected to hear and what they actually heard. McGuire et al. concluded that the VSM system resides primarily in the lateral temporal cortex bilaterally. This cortical region has been shown to be differentially activated by different tasks involving elicited verbal mentation (McGuire, Silbersweig, Murray, et al., 1996). Whereas silent reading does not activate this region, it is activated by participants' imagining speech in another person's voice (Shergill et al., 2001). While it seems plausible that generating verbal mentation in one's own voice will be a relatively automatic process, mentally imitating another voice presumably requires internal inspection of this imagined speech (necessary to assess whether the voice has the prosody, tone, pitch, and rhythms of the voice it is intended to be; McGuire, Silbersweig, Murray, et al., 1996), placing high demands on the VSM system. Greater demands of mentally imitating another voice are consistent with healthy participants rating such a task as subjectively more difficult than silent articulation (Shergill et al., 2001). The implication of the lateral temporal cortex in VSM accords with findings that the temporal lobe is associated with speech perception (Wise et al., 1991; Zatorre, Evans, Meyer, & Gjedde, 1992) and the phonological processing of heard words (Demonet et al., 1992). Furthermore, patients with left temporal lesions have been found to have impairments in imagining others speak (Zatorre & Haplern, 1993).

In addition to imaging studies, other techniques, such as EEG studies (e.g., Ford & Mathalon, 2005), have been employed to examine the neurophysiological signaling (corollary discharge) mechanisms involved in VSM. These techniques have also been applied to the study of individuals assumed to have a deficit in VSM, such as those who experience AVHs, with the result that a wide range of neuroimaging and neurophysiological studies with normal and AVH-hearing populations (primarily those with a diagnosis of schizophrenia) have been published. In the first part of this article, we review these findings and evaluate their significance for any comprehensive theory of AVHs.

1.1. Neuroimaging of silent articulation

A number of neuroimaging studies have been performed to investigate the neural correlates of silent articulation in healthy participants (e.g., Bullmore et al., 2000; Friedman et al., 1998; McGuire, Silbersweig, Murray, et al., 1996; Shergill et al., 2001). In Shergill et al.'s (2001) fMRI study, participants listened to single words, presented via

headphones, read out by a neutral voice at 3 s intervals. In the baseline condition subjects were simply instructed to listen to each word carefully. In what was termed the Inner Speech condition, participants were asked silently to complete sentences of the form 'I like *x*' or 'I like being *y*' using the presented word. Functional MRI scans of brain activation during the Inner Speech condition were compared to baseline, and the Inner Speech condition was found to be associated with left-sided activation, in Broca's area (BA 44), Wernicke's area (BA 22), the SMA (BA 6), and the insula as well as the superior parietal lobe (BA 7) and right posterior cerebellar cortex. These findings accord with those of an earlier fMRI study by Bullmore et al. (2000), which found similar neural regions to be specifically associated with inner speech. In this experiment, participants were presented with a word, such as 'goat', during scanning, and then had to decide whether it was a living or nonliving object. This decision then had to be articulated in inner speech. However, rather than just examining the neural correlates of this inner speech compared to a control condition (fixating on a blank screen), Bullmore et al. (2000) attempted to link the activation they found in specific neural regions to the function that it was associated with. This led them to propose that activation in a series of connections from the prefrontal cortex to the left inferior frontal gyrus/Broca's area (BA 44, 45) via the SMA (BA 6) was responsible for subvocal planning and articulation. Inferior parietal lobule (BA 39, 40) activation extending to the precuneus (BA 7) and the superior temporal gyrus (BA 21) was presumed to represent the monitoring of the subvocal output of Broca's area.

The question of whether AVH-hearers have different patterns of brain activation during silent articulation as compared to healthy controls was first broached by McGuire et al. (1995). The psychiatric sample in this study consisted of patients diagnosed with schizophrenia who had consistently experienced AVHs during their illness, but who were in remission and had low levels of psychotic symptoms at the time of testing. PET scans were performed on both healthy participants and these remitting AVH-hearers as they silently read single words presented on a visual display unit. Analysis of the results found no differences in the areas activated by AVH-hearers and controls during silent articulation. This result was replicated by Shergill, Bullmore, Simmons, Murray, and McGuire (2000), using the same paradigm as Shergill et al. (2001) described above. In one part of Shergill, Bullmore, et al.'s (2000) study, the presented words were used to complete sentences such as 'I like *x*', or 'I like being *y*', which were silently articulated by the participants. fMRI imaging was used to compare the neural correlates of this silent articulation between healthy controls and AVH-hearers. It was found that patients showed no differences to controls in neurological activation when performing this task.

Such findings suggest that silent recitation of sentences is not abnormal in AVH-hearers in remission. Because silent articulation is believed to result in low levels of activation of the VSM system (McGuire, Silbersweig, Murray, et al., 1996), this suggests that tasks involving low levels of VSM are not performed abnormally by AVH-hearers. However, Shergill et al. (2003) hypothesized that, if the VSM load involved in silent articulation could be increased, then differences would be found between AVH-hearers and healthy controls. Accordingly, Shergill et al. (2003) asked participants to covertly articulate the word 'rest' either once every 4 s (low-VSM condition) or once every second (high-VSM condition). It was found that lateral temporal cortex activation (signifying higher levels of VSM) increased with the faster rate of covert generation in healthy participants. In contrast, in patients with a history of AVHs (but, at the time of study, in clinical remission) less activation was shown in the right (but not left) superior temporal gyrus, the right parahippocampal gyrus, and the right cerebellar cortex when compared to controls performing the faster rate of covert generation.

In conclusion, evidence suggests that neural correlates of inner speech during silent recitation of sentences are not abnormal in AVH-hearers in remission. Because silent articulation is thought to involve low levels of activation of the VSM system (McGuire, Silbersweig, Murray, et al., 1996), it seems likely that any neural differences between AVH-hearers and healthy controls will remain obscured unless methods other than low-VSM tasks are employed.

1.2. Neuroimaging of first-, second-, and third-person inner speech: auditory verbal imagery

In addition to studies of silent articulation of sentences, a number of studies have investigated the neural correlates of inner speech in the form of imagined speech, termed auditory verbal imagery (AVI). Such a study was performed by Shergill et al. (2001) with healthy participants, using the paradigm of completing sentences such as 'I like *x*', or 'I like being *y*', as outlined above. However, in addition to silently articulating such sentences, participants were also asked to imagine either the sentence being spoken in their own voice (first-person AVI), or sentences of the form 'You like *x*', being spoken to them (second-person AVI), or sentences of the form 'He likes *x*', as if spoken about them (third-person AVI). When imagining the sentences being spoken to them or about them, participants were instructed to imagine the

voice as being the same voice that spoke the words that were played to them. Shergill et al.'s (2001) fMRI analysis found that auditory verbal imagery in general, as indexed by the three imagery tasks combined, was associated with activation typically found in 'inner speech' tasks (left inferior frontal and parietal cortex and SMA). In addition, activation was seen in the left precentral and superior temporal gyri, and the right homologues of all these areas. The involvement of the right hemisphere homologues was presumed to occur due to retrieving and processing the prosody of the imagined other's speech.

McGuire et al. (1995) were the first to address the question whether, when producing auditory verbal imagery, patients with schizophrenia who have experienced AVHs have a different pattern of activation to healthy controls. Their study required participants to listen to recorded single words and then to imagine sentences of the form 'You are *x*', or 'You are a *y*', being spoken to them by the voice on the recording (this study only addressed second-person AVI). Using the results of their PET scans, it was found that AVH-hearers in remission showed reduced activation in the left middle temporal gyrus (MTG), the rostral SMA, and the left medial prefrontal cortex compared to healthy controls. It was also found that AVH-hearers in remission, compared to patients with schizophrenia who did not experience AVHs, showed reduced activation in the left MTG, the rostral SMA, and a posterior region spanning primary and secondary visual areas and the adjacent cerebellar cortex.

A difference in neural activation during AVI between patients with schizophrenia who had experienced AVHs and healthy controls was also found by Shergill, Bullmore, et al. (2000). As reviewed above, participants were asked silently to articulate sentences of the form 'I like *x*', or 'I like being *y*', and additionally to generate first-, second-, and third-person AVI. Functional MRI analysis of the AVI conditions, using an activation map for all three types of AVI combined, showed differences between patients with schizophrenia who had heard voices (but who were in remission) and healthy controls. Specifically, when AVH-hearers in remission imagined others speaking to them, there was less activation in the posterior cerebellum, hippocampal complex, and lenticular nuclei bilaterally, and also the right thalamus, middle and superior temporal gyri, and left nucleus accumbens.

In conclusion, evidence suggests that mechanisms underlying the ability to silently imagine a voice speaking are atypical in AVH-hearers in remission. Because imagining others speak is thought to be associated with high levels of activation of the VSM system (McGuire, Silbersweig, Murray, et al., 1996; Shergill et al., 2001), these conclusions are consistent with the findings of Shergill et al. (2003) that neural activation on tasks involving high levels of VSM (in this instance, producing inner speech at a high rate) may be abnormal in AVH-hearers in remission.

1.3. Event-related potential studies of silent articulation

Whereas scanning techniques such as PET and fMRI are able to provide indirect evidence that VSM may be atypical in individuals affected by AVHs, studies using event-related potentials (ERPs) are able more directly to test this hypothesis. As noted earlier, the VSM system is thought to work through a corollary discharge system whereby discharges from the motor speech-producing areas in the frontal lobes 'warn' the auditory cortex that self-produced speech is about to be produced (Creutzfeldt et al., 1989). This is postulated to work through the motor speech-producing areas' corollary discharge deactivating the auditory cortex when self-produced speech is about to occur (Ford & Mathalon, 2005).

Ford et al. have used ERPs to study the activation of such areas during inner speech in patients diagnosed with schizophrenia (Ford et al., 2001). In this experiment, N1 ERPs were used as a measure of engagement of the auditory cortex during inner speech. N1 is generated by the superior temporal gyrus of the auditory cortex in response to auditory stimuli of all kinds (ibid). Participants were asked to repeat typical hallucinatory statements (e.g., 'That was really stupid') silently to themselves for 30 s, while brief auditory stimuli were presented via headphones. The instruction to participants to use silent articulation in this study would presumably have entailed a low load on the VSM system (as opposed to the high load that would have resulted from instructions to use AVI). It was found that, in healthy participants, the N1 response to brief auditory sounds during 'inner speech' (construed here as repeating silently the hallucinatory type statements) decreased significantly as compared to a baseline condition involving silently fixating on a point. In contrast, patients with schizophrenia showed no significant difference in N1 response between these two conditions. Ford et al. interpreted this to show that, whereas healthy individuals dampen their auditory cortex response during self-produced speech, as if to warn the auditory cortex that the sound is internally produced, patients with schizophrenia do not. As a consequence this may lead patients with schizophrenia to experience their own self-produced speech as having an external source. However, Ford et al. failed to find a significant correlation between levels of hallucinations (as per SAPS summary score for hallucinations) and the N1 effect (baseline less 'inner speech'),

suggesting that other factors must be causally involved in AVHs beyond damping of the auditory cortex response during the silent recitation of sentences.

No electrophysiological studies have been performed to examine directly whether corollary discharge is abnormal in AVH-hearers during high-VSM inner speech tasks. One study that used overt speech rather than inner speech attempted to examine whether corollary discharge was abnormal in AVH-hearing patients with schizophrenia (Ford & Mathalon, 2005, Section 6). These researchers assessed the involvement of frontal lobe speech production areas with speech reception areas in the temporal lobe, using measures of EEG coherence (frontal–temporal gamma synchrony). In the Talking condition participants uttered aloud the syllable [a] every 1–2 s. They then heard playback of their speech through earphones at various degrees of distortion (no distortion, half a semitone lower, and a full semitone lower). In the Listening condition, they did not speak but simply heard undistorted pre-recorded sounds from the Talking condition. The difference in coherence between the Talking and Listening conditions (a measure of the interdependence of the frontal and temporal lobes during talking), although different in participants with schizophrenia when compared to normal controls, did not correlate with hallucination severity (Ford & Mathalon, 2005, Section 8). Furthermore, the difference in coherence between the distorted and undistorted Talking conditions also did not correlate with hallucination severity. However, this study used overt speech rather than inner speech, and there is persuasive neuroimaging evidence that different neural regions (aside from motor areas) are activated in covert and covert speech (Huang et al., 2001; Soltysik & Hyde, 2006). As Ford and Mathalon (2005) note, "talking may be a poor proxy for thoughts and voices" (p. 188).

From this series of experiments Ford and Mathalon (2005) concluded that "we have not been able to demonstrate a clear relationship between neurobiological indicators of dysfunctional corollary discharge and the extent to which a patient currently reports experiencing auditory verbal hallucinations" (p. 187). Thus, ERP studies provide good evidence that low-VSM inner speech tasks do not appear to result in differential neural activation specific to the experience of AVHs. It still remains for ERP studies to examine neurophysiological differences between AVH-hearers and controls when performing high-VSM inner speech tasks such as AVI. We would predict that, on the basis of the neuroimaging findings reviewed above, hallucination severity would correlate with the N1 effect when the participants are imagining others speak.

1.4. Summing up

The neuroimaging and neurophysiological findings reviewed above provide evidence that the neural correlates of inner speech, conceptualized as the silent articulation of sentences, are not fundamentally different in AVH-hearers in remission compared to healthy controls. The neuroimaging studies also show that, when one's own self or others are imagined talking in inner speech, there is a neural difference between AVH-hearers in remission and healthy controls without AVHs. As noted above, the silent articulation of sentences places a low load on the VSM system, whereas repeating single words at an increased pace or imagining a sentence being spoken place a high load on the VSM system. This leads to the conclusion that predisposition to AVHs is associated with abnormal neurological activation during tasks that involve high-VSM loads.

An important caveat to make at this point is that the neuroimaging evidence showing abnormal neurological activation in remitting AVH-hearers during high-VSM inner speech tasks does not necessarily translate into different patterns of judgment about the authorship of externally heard voices. Using a well-established paradigm (Johns, Gregg, Vythelingum, & McGuire, 2003), Johns, Gregg, Allen, and McGuire (2006) examined such judgments in patients with schizophrenia with and without current AVHs. Participants spoke into a microphone and then had the sound played back to them over headphones in real time. The voice they heard played back to them could be their own voice, their own distorted voice, another person's voice (saying the same word), or another person's distorted voice. The study found that patients who had previously experienced AVHs, but who were currently hallucination-free, did not misattribute the source of the spoken voice any more than healthy controls. Thus, although the neurological activation on high-VSM tasks differs between such patients compared to healthy controls, this does not appear to result in a detectable difference in tasks that require judgments about the authorship of heard voices. In contrast, Johns et al. (2006) found that patients currently experiencing regular AVHs were more likely (compared both to controls and to remitting AVH-hearers) to misattribute the source of the heard voices.

There are at least two possible ways of accounting for this apparent discrepancy between the neural correlates findings and the work of Johns et al. (2006) on judgments of voice authorship. First, it may be that both remitting and current AVH-

hearers show the same patterns of activation on high-VSM tasks, but that the deviation from typical activation on such tasks is greater in current AVH-hearers. The second possibility is that other factors, coupled with this basic neurological abnormality on high-VSM tasks, distinguish current from remitting AVH-hearers. With regard to the first possibility, one point that can immediately be made is that equivalents to the neuroimaging studies reviewed above have not yet been performed with current (as opposed to remitting) AVH-hearers. Other methodologies, such as Ford and Mathalon's (2005) use of neurophysiological indicators on an overt (rather than inner) speech task, were not able to demonstrate a relation between the neural activation of the VSM system (dysfunctional corollary discharge) and the extent to which a patient currently reports experiencing AVHs. This suggests that it is not simply that current AVH-hearers show greater deviation from typical patterns of response to high-VSM tasks compared to AVH-hearers in remission.

At present, the balance of evidence seems to suggest that, if impairment in VSM does play a role in causing AVHs, it must do so in conjunction with other factors. One potential factor is the activation level of the right anterior cingulate gyrus, previously found to be implicated in cognitive tasks involving mental effort (Bush, Luu, & Posner, 2000) including monitoring the predicted consequences of actions (Ito, Stuphorn, Brown, & Schall, 2003) and the regulation of consciously detected conflicts (Dehaene et al., 2003). Anterior cingulate activity has been found to correlate with measures of AVHs (Silbersweig et al., 1995), and to be activated during AVHs (Lennox, Bert, Park, Jones, & Morris, 1999; Shergill, Brammer, et al., 2000). In addition, Hubl et al. (2004) found that white matter left cingulate bundle activation differentiated between auditory hallucinators and non-hallucinators. The anterior cingulate gyrus would therefore appear to be a plausible candidate for a factor that interacts with VSM deficits in the genesis of full-blown AVHs.

In sum, research on the neural correlates of AVHs supports the contention that VSM of inner speech is atypical in AVH-hearers, but that this difference only obtains on tasks involving high-VSM demands, such as auditory verbal imagery (AVI). In contrast, paradigms involving low-VSM demands, such as silent articulation, do not distinguish (neurologically, at least) AVH-hearers from healthy controls. And yet it is precisely these paradigms that are commonly held to be the best experimental analogues of normal inner speech (e.g., McGuire et al., 1995). This leads us to a conundrum: if we wish to account for the genesis of AVHs in terms of neurally instantiated VSM deficits leading to the impaired self-monitoring of inner speech, this cannot involve the kinds of verbal mentation (low-VSM tasks such as silent articulation) that are typically held to be the best experimental approximations to typical inner speech. Specifically, it needs to be explained why AVH-hearers might be performing verbal mentation in inner speech associated with high levels of VSM. Why, during typical inner speech, should they be performing tasks utilizing the same cognitive resources that are involved with imagining people speak? We suggest that this conundrum can be resolved through a careful examination of the nature, forms, and developmental significance of inner speech.

2. What is this thing called inner speech?

We propose that at least some of the apparent contradictions that have emerged from empirical research on inner speech stem from two related sets of unexamined assumptions: First, about what inner speech is, and second, about how it can be elicited artificially. In this section, we examine these assumptions in light of contemporary empirical and theoretical work on inner speech and its developmental precursors.

Many neuroimaging studies of AVHs have drawn on a conception of inner speech which derives from Levine, Calvanio, and Popovics (1982). In a paper examining the possibility of language in the absence of inner speech, Levine et al. defined inner speech as the "subjective phenomenon of talking to oneself, of developing an auditory–articulatory image of speech without uttering a sound" (p. 391). Some studies explicitly refer to this definition (e.g. Evans, McGuire, & David, 2000), while others assume it implicitly by holding the silent articulation of sentences to represent inner speech (e.g. Shergill et al., 2001). Other studies have initially defined inner speech as "thinking in words" (McGuire et al., 1995, p. 596), while going on to operationalize inner speech for experimental purposes as the subvocal recitation of predefined sentences. There has been a certain degree of consistency in the equation of inner speech with subvocal recitation. By these definitions, patients and controls are said to be performing inner speech when they mentally recite sentences such as "You are stupid" (McGuire et al., 1995, p. 597).

Such a subvocal rehearsal/recitation conception of inner speech has a clear historical foundation in (Baddeley and Hitch's, 1974) model of working memory (Baddeley, 1986). This model proposed the existence of an articulatory loop consisting of a phonological store, capable of holding speech-based information, and an articulatory control process. In Baddeley and Hitch's model, the articulatory loop is responsible for the short-term maintenance of up to two seconds'

worth of verbal information in working memory. It thus corresponds to the part of the phonological information-processing system that Baddeley (1986) described as the "voice in the head", or inner speech. The large body of empirical support for Baddeley and Hitch's working memory model (e.g., Baddeley, 2001) may be one reason why the subvocal rehearsal/recitation conception of inner speech has influenced experimental attempts to elicit inner speech in neuroimaging studies. A second, related reason is that the elicitation methodologies that derive from this conception are relatively straightforward and thus appropriate for the considerable practical constraints associated with neuroimaging studies.

Baddeley and Hitch's conception of inner speech as resulting from activation of the articulatory loop is not, however, the only conception of inner speech to have proved valuable to cognitive scientists. In Vygotsky's (1934/1987) theory, inner speech represents the endpoint of a developmental process in which external discourse gradually becomes internalized to form verbal thought. Generally speaking, Vygotsky's hypotheses about the development of inner speech (and its semi-covert precursor, private speech) have been supported by empirical research (Berk, 1992; Winsler, 2004). Of particular interest for our purposes is the finding that private speech frequently reflects the dialogic structure of interpersonal verbal exchanges (e.g., Ramirez, 1992; Wertsch, 1980).

Vygotsky's (1934/1987) ideas about inner speech form part of a broader theory of the social origins of higher mental processes. This theory is encapsulated in his well-known claim that every mental function appears twice in development: First on the interpsychological plane, as a function distributed between individuals, and second on the intrapsychological plane, as an internalized version of that previously external function (Vygotsky, 1931/1997). In the case of the developmental transition from external speech to inner speech, this internalization can be seen at work, for example, in a mother and child solving a jigsaw puzzle. Initially, the puzzle-solving process will involve an external dialogue between mother and child, with the mother typically asking the child questions about which piece should be placed where, the child answering, and so on. Later on this dialogic pattern of exchange is internalized as inner speech (Fernyhough, 2004). Mental dialogue is, therefore, an internal version of the interplay of perspectives that takes place between individuals on the external plane (Fernyhough, 1996).

For Vygotsky, internalization of previously external verbal activity was an essential component in children's developing self-regulation of behavior and cognition. Accordingly, studies of children's private speech have shown that it often contains utterances that have a self-directive function (Luria, 1961). The dialogic nature of private speech (and, by extension, inner speech) guarantees that children become able to take the role of the questioner, advisor, and director in the regulation of their activity. As Vygotsky (1930–1935/1978) put it, the use of verbal mediation means that humans are able to "control their behavior from the outside" (p. 40).

Central to the Vygotskian conception of inner speech is the assumption that, like its developmental precursor private speech, it will retain certain characteristics of the external discourse from which it is derived. Several authors (e.g., Wertsch, 1980) have noted that one implication of Vygotsky's theory is that inner speech will have a dialogic structure. In arguing against a view of inner speech as a homogeneous phenomenon, one of us (Fernyhough, 2004; submitted for publication) has previously suggested that at least two distinct forms of dialogic inner speech might be predicted on the basis of Vygotsky's theory. These are *expanded inner speech*, where the flow of verbal mentation retains the give-and-take quality of external dialogue, and *condensed inner speech*, where the linguistic accoutrements of dialogue become jettisoned and inner speech becomes a process of "thinking in pure meanings" (Vygotsky, 1934/1987).

Fernyhough's (2004) four-stage model of the development of inner speech is represented diagrammatically in Fig. 1. At Level 1 (external dialogue), the child and caregiver engage in verbally mediated reciprocal exchanges such as might arise in the collaborative solving of a jigsaw puzzle. At Level 2 (private speech), children conduct these dialogues in overt self-directed speech. This linguistic activity gradually becomes internalized, leading to the creation of expanded dialogic inner speech (Level 3). Finally, the syntactic and semantic abbreviation processes described by Vygotsky effect its transformation into condensed dialogic inner speech (Level 4), corresponding to Vygotsky's stage of "thinking in pure meanings".

This account entails that thinking retains the dialogic, semiotically mediated qualities of interpersonal discourse even when, phenomenologically speaking, it does not share the give-and-take character of external dialogue. To put it another way, the conversation that we have with ourselves does not cease when it ceases to be conducted in explicit, syntactically expanded questions and answers. One of the present authors (Fernyhough, 1996, 2004, 2005, submitted for publication) has previously examined the implications of a dialogic view of the higher mental processes for our understanding of a range of different cognitive processes, including executive functioning and social understanding. As Carruthers (2002) notes, most members of the cognitive science community endorse a purely communicative conception of language, which sees language as an input–output system for central cognition, with thinking undertaken

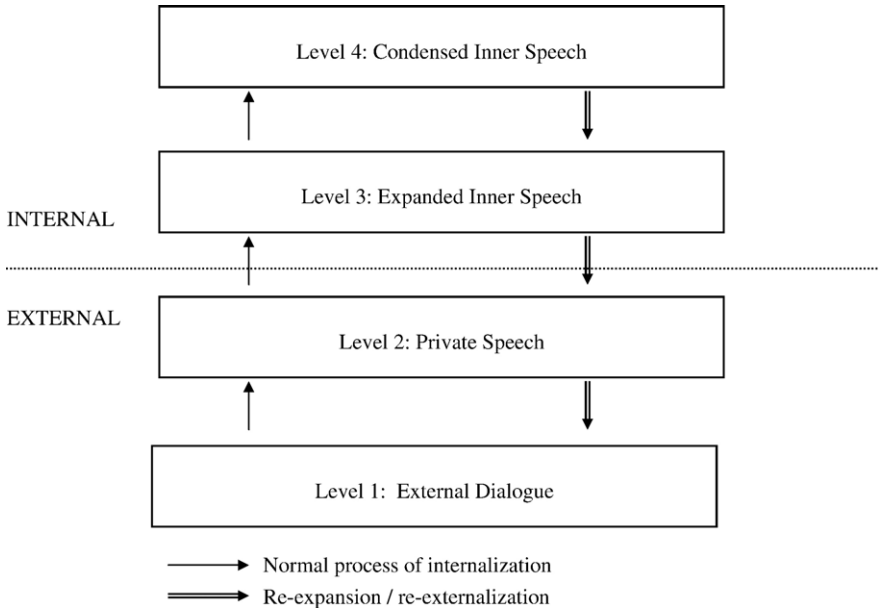

Fig. 1. Fernyhough's (2004) four-stage model of the development of inner speech.

in some other form of representation. For example, Fodor (1975) proposed that we think, not in natural language, but in what is commonly termed mentalese. At the other end of the spectrum exist views, such as those held by early behaviorists, which equate thought with inner speech, claiming thought to be merely speech without sound (e.g., Watson, 1920). Treading a middle path between these two poles, a Vygotskian account of inner speech fits with cognitive conceptions of language which see it as having a constitutive role in cognition that goes beyond the straightforward communication of information (e.g., Carruthers, 2002). Vygotsky (1934/1987) proposed that thought and speech could be visualized as two intersecting circles. The overlapping region of thought and speech represents what may be termed verbal thought or inner speech. Such a conception allows that, while thought may occur in the medium of inner speech, nonetheless "[t]here is a large range of thinking that has no direct relationship to verbal thinking" (Vygotsky, 1934/1987, p. 115). This position is consistent with evidence of the need for verbal thought in some forms of cognition (Hermer-Vazquez, Spelke, & Katsnelson, 1999), as well as the possibility of thought without language (Weiskrantz, 1988). In line with this conclusion, an elegant neuroimaging experiment performed by Dehaene, Spelke, Pinel, Stanescu, and Tsivkin (1999) has demonstrated the involvement of language in performing exact mathematical calculations in one's head, but not in the performance of approximate arithmetic. We thus adopt a Vygotskian position that inner speech (both expanded and condensed versions) forms a subset of human thought processes.

We argue that Fernyhough's (2004) extension of Vygotsky's conception of inner speech has at least two important implications for our understanding of AVHs. First, the distinction between condensed and expanded inner speech allows us to understand the dynamic relation between AVHs, where inner speech is experienced with its full complement of acoustic properties, and much of regular human mentation, where the acoustic properties of the voices in inner dialogue are less apparent. In expanded inner dialogue, which we suggest forms the raw material of AVHs, human thought incorporates a multiplicity of internalized voices. If such expanded dialogues do indeed form the raw material of AVHs, it is unsurprising that the existing thoughts and ideas of voice-hearers may come to be reflected in part or much of the content of the AVHs (Leudar & Thomas, 2000). In terms of the neural correlates evidence discussed above, the dialogic conception of inner speech consequently allows us to see how different forms of inner speech entail different levels of VSM demands. Specifically, conducting expanded inner dialogues will require high levels of VSM, while condensed inner speech (where the acoustic properties of inner speech are attenuated) does not make such demands on the VSM system.

A second implication of this conception of inner speech is that it predicts the characteristic pragmatic qualities of AVHs. Specifically, it allows us to understand how AVHs frequently take the form of commands. As inner speech is

developmentally linked with the control of action (Luria, 1961; Vygotsky, 1934/1987), we should not be surprised that AVHs frequently have a similar regulatory quality.

3. Neural correlates of inner dialogue: testing a Vygotskian model

We turn now to the question of how the paradigms used to elicit inner speech in neuroimaging and neurophysiological studies relate to this alternative, dialogic conception of inner speech. Recall that the majority of such studies have elicited inner speech by asking participants to repeat sentences (such as 'I like swimming') subvocally. At first glance the resulting subvocal speech would appear closer in form to Level 3 (expanded) inner speech than Level 4 (condensed) inner speech. Crucially, though, the dialogic quality of typical spontaneous inner speech is absent in such experimentally elicited utterances. If our Vygotskian characterization of inner speech is accurate, we would have to conclude that the methods used for eliciting subvocal speech in typical neuroimaging studies do not lead to the generation of anything resembling naturally occurring spontaneous inner speech.

Furthermore, Vygotsky's ideas about the semiotic mediation of higher mental processes entail that verbal mentation should continue even when inner speech loses its explicit linguistic form. Specifically, Fernyhough's (2004) extension of Vygotsky's ideas on this issue suggest that, with the exception of episodes when inner speech is temporarily expanded to form Level 3 (expanded) inner speech, verbal mentation should consist of ongoing Level 4 (condensed) inner speech. The implications of this for neuroimaging studies are, we suggest, profound. In a study such as that of Shergill et al. (2001), the neural correlates of inner speech are examined by comparing a baseline condition (subjects listening to single words presented at 3 s intervals) to a condition where they are silently articulating sentences. However, what such studies treat as a baseline condition is likely to involve ongoing condensed inner speech. The results of such studies would therefore appear to be contaminated by the persistence of a form of inner speech even into the baseline condition.

The dialogic conception of inner speech has value, however, beyond pointing up possible methodological limitations of existing neural correlates studies. One of the present authors (Fernyhough, 2004) has put forward a theory of AVHs that draws on Vygotskian ideas about the developmental significance of inner speech. According to this hypothesis, AVHs result from the temporary re-expansion of Level 4 (condensed) inner speech into Level 3 (expanded) inner speech, particularly under conditions of stress and cognitive challenge. Such movement between forms of inner speech is held to be a typical feature of human mental life. In individuals who report AVHs, in contrast, the re-expansion of condensed inner speech interacts with pre-existing source monitoring and other cognitive biases to result in the expanded dialogue being experienced as involving an alien voice. Fernyhough (2004) further proposed that such hallucinations in psychiatric patients may be shaped (in terms of their content and structure) by the experience of trauma.

Fernyhough (2004) suggested that the re-expansion model of AVHs could be tested (and distinguished from other possible hypotheses drawn from a Vygotskian account of inner speech) in four main ways. First, the re-expansion model entails that AVH-hearers should experience normal Level 4 (condensed) inner speech. To date, this issue has not been investigated empirically. Second, the re-expansion model would predict that AVH-hearers should not experience normal Level 3 (expanded) inner speech, as such speech would typically be experienced as having an alien quality. Again, this prediction awaits empirical examination. Third, the model would predict that AVHs would be associated with stress and cognitive challenge, a finding that receives support from a range of sources (e.g., Cooklin, Sturgeon, & Leff, 1983). Finally, the model holds that previously unaffected individuals should occasionally experience AVHs under conditions of very great stress, a prediction that is supported by reports of auditory hallucinations following bereavement and extreme stress (Balan et al., 1996; Reese, 1971).

Much more empirical research needs to be done before the re-expansion model of AVHs can be properly evaluated. In the meantime, several resulting hypotheses about the neural correlates of inner speech and AVHs can be assessed in light of the neuroimaging and neurophysiological evidence reviewed above.

Hypothesis 1. Neural correlates of condensed inner speech in AVH-hearers will be identical to those in healthy controls.

Fernyhough's (2004) re-expansion model holds that AVH-hearers should experience normal condensed inner speech. The prediction can thus be made that the neural correlates of this speech should not differ from those observed in healthy controls. One way of making sense of this prediction is in terms of our earlier distinction between tasks

involving low and high levels of VSM demands. The attenuation of the acoustic characteristics of speech in Level 4 (condensed) inner speech is, we suggest, likely to result in a reduced VSM load in inner speech of this kind. One would therefore expect to see Broca's area activation accompanied by only minor corresponding activation in the VSM system (i.e., bilaterally in the lateral temporal cortex) in both AVH-hearers and non-AVH-hearers. The testing of this hypothesis would require a task that elicits condensed inner speech, with no expanded (Level 3) component. Currently no ecologically valid procedures exist for examining condensed inner speech, although the procedure employed by Ford et al. (2001) might be construed as offering an approximation to such a condition. Recall that, in Ford et al.'s baseline condition, participants had to fixate on a screen, while auditory stimuli (speech sounds, noises, etc.) were played to them. As Ford et al. themselves note, "spontaneous inner speech" (p. 1914) could occur during this baseline condition. From a Vygotskian perspective, we would expect condensed (Level 4) inner speech to be occurring during this part of the task. Thus, we would predict no N1 activation differences between AVH-hearers and healthy controls during this condition. In line with our prediction, Ford et al. found that N1 responses during the baseline condition, although being slightly (but not significantly) lower in patients than in controls, were not related to the presence of AVHs.

Hypothesis 2. Neural correlates of inner speech in AVH-hearers will differ from those of controls in tasks involving the re-expansion of condensed inner speech.

The central proposal of the re-expansion model is that AVHs stem from the re-expansion of condensed inner speech. AVH-hearers are suggested to differ from their unaffected counterparts in their experience of Level 3, but not Level 4, inner speech. To date, no neural correlates studies have examined the transition between these forms of inner speech. The paradigm that would lend itself most easily to being adapted to such an investigation is that of Bullmore et al. (2000). In this study, participants viewed a word on a screen, then had to articulate internally a semantic judgment about the stimulus. We speculate that the semantic judgment process might involve Level 4 (condensed) inner speech, with a transition to Level 3 (expanded) inner speech during the subvocal articulation of the judgment. In line with our first hypothesis, we would predict that the neural activation of AVH-hearers would not differ from controls up to the point when condensed inner speech stops and the conclusion becomes subvocally articulated. In contrast, the transition to Level 3 inner speech during the subvocal articulation of the semantic judgment should lead to observable differences between AVH-hearers and controls. In addition to typical patterns of activation associated with high-VSM tasks (corresponding here to the occurrence of Level 3 inner speech), one might also expect AVH-hearers to show greater levels of parietal cortex activation, representing the experience of alienness (Spence et al., 1997). Time-sequenced fMRI studies might be best placed to detect any such group differences in neural correlates of inner speech in the transition between Level 3 and Level 4 inner speech.

Hypothesis 3. Neural correlates of elicited expanded inner speech should differ between AVH-hearers and controls.

As noted above, the re-expansion model entails that AVH-hearers should not experience typical Level 3 inner speech. To date, this prediction has been tested neither by examination of subjective reports of AVH-hearers' inner speech, nor by the attempted experimental elicitation of expanded inner speech. We have noted that no existing neural correlates studies have included explicit instructions to generate expanded dialogic inner speech. The closest approximations in the existing literature would appear to be the studies of McGuire et al. (1995) and Shergill, Bullmore, et al. (2000), where second-person AVI has indeed been shown to be associated with neural differences between AVH-hearers and controls. In contrast, no such group differences are found in the neural correlates of silent (non-dialogic) articulation. Our hypothesis here could be tested simply by asking participants to conduct an expanded dialogue in their heads while undergoing imaging. We would predict that such a condition would lead to similar differences in neural activation to those predicted by Hypothesis 2. We would also predict, in line with the re-expansion model, that the experience of expanded inner dialogues, even when experimentally elicited, should, subjectively, have some of the hallucinatory quality of the AVHs held by the re-expansion model to result from the re-expansion of condensed inner speech.

Hypothesis 4. Neural correlates of the experience of an AVH should be similar to patterns of activation in healthy individuals undertaking expanded inner speech.

We noted above (Hypothesis 3) that approximations to the elicitation of expanded inner speech exist in studies that have required participants to undertake AVI. A pertinent observation in this respect is one made by Shergill, Brammer, et al. (2000). Reporting their use of fMRI scanning of patients with schizophrenia while they were experiencing AVHs,

Shergill, Brammer, et al. commented that "The pattern of activation we observed during auditory hallucinations is remarkably similar to that seen when healthy volunteers imagine another person talking to them (auditory verbal imagery)" (p. 1036). Specifically, they observed common activation of the bilateral frontal and temporal gyri, along with right-sided precentral and inferior parietal gyri. One important difference between patients hearing AVHs and healthy participants generating AVI, however, is that the latter is associated with increased SMA activation, while this region is only weakly activated during AVHs. We can therefore conclude that Hypothesis 4 receives some support from studies that involve approximations to expanded inner speech (through the elicitation of AVI in healthy controls), but that the involvement of the SMA may differ between AVH-hearing patients and healthy controls generating AVI.

Finally, we suggest that the re-expansion hypothesis may be able to resolve the controversy about the involvement of the inferior frontal gyrus (Broca's area) in AVHs. While a number of studies have found increased left inferior frontal cortex activation (compared to baseline) in AVH-hearers experiencing hallucinations (Shergill et al., 2004; Shergill, Brammer, et al., 2000), others have failed to find such an increase in activation during AVHs (Lennox et al., 2000; Silbersweig et al., 1995). These negative findings on the involvement of Broca's area activation in AVHs have been interpreted as providing "no support to the theory that auditory hallucinations arise from abnormalities of inner speech" (Lennox et al., 2000, p. 19).

We would argue that these contradictory findings may be accounted for by the persistence of Level 4 (condensed) inner speech into baseline conditions in typical neuroimaging studies. Given that the re-expansion model predicts that AVH-hearers will experience normal Level 4 (condensed) inner speech, this would result in continued activation of Broca's area even when patients are being scanned at rest, without any current experience of hallucinations. Broca's area activation will also be a feature of the experience of hallucinations, held by the re-expansion model to involve Level 3 (expanded) inner speech. Thus, the difference in Broca's area activation between the AVH and non-AVH states will depend on how much condensed inner speech patients have been performing during their non-hallucinating rest period in the scanner. If patients are undertaking continuous condensed inner speech at rest, this would be likely to lead to no difference in Broca's area activation between AVH and non-AVH phases. In contrast, if low levels of condensed inner speech are being performed at rest, then an increase in Broca's area activation is likely to accompany AVHs. In order to resolve this question, we recommend that future neuroimaging studies incorporate post-scanning self-report measures of the thinking and inner speech that occurred during participants' time in the scanner. If such subjective reports are taken into account, it should be possible to determine the relative contributions of condensed and expanded inner speech to Broca's area activation in both hallucination and rest phases of such imaging studies.

4. Conclusions

At the start of this paper we set out certain key features of AVHs that we believe a comprehensive account of the phenomenon must explain. We reviewed the neuroimaging and neurophysiological evidence that has to date been used to study neural correlates of inner speech in AVH-hearers and healthy controls. We proposed that an apparent conundrum posed by this research, namely that differences between these groups only obtain on tasks involving high levels of VSM, can be resolved by considering the different possible forms that inner speech can take. Fernyhough's (2004) re-expansion model of AVHs was used to generate testable hypotheses for future studies of the neural correlates of inner speech in both AVH-hearers and healthy controls.

Before drawing any final conclusions about the value of the theoretical ideas considered here, we return to our earlier caveat about the assumed, but not proven, involvement of inner speech in AVHs. It has previously been argued that research showing that neither the articulatory loop (David & Lucas, 1993; Haddock, Slade, Prasaad, & Bentall, 1996) nor the inner-voice/inner-ear partnership is impaired in patients with schizophrenia means that, if any relationship exists between AVHs and inner speech, it is not simplistic or direct (Evans et al., 2000). While this conclusion appears sensible, we note that it assumes that working memory conceptions of inner speech (as silent articulation/rehearsal) map precisely onto those derived from Vygotsky's theory. Rather, we have argued that spontaneous inner speech has varieties of forms, as well as syntactic and semantic properties, that make it a richer phenomenon, both in its subjective qualities and its neural instantiation, than the subvocal articulation of external utterances. Lack of any strong evidence for articulatory loop deficits in schizophrenia cannot therefore be taken to invalidate the re-expansion model, as both condensed and expanded inner speech will be reliant on a functioning phonological apparatus.

We now return to the four key challenges for any comprehensive account of AVHs. The first task is to explain why AVH-hearers experience a voice in the absence of any external stimulus. Our answer to this question, in line with those

of many other authors (e.g., Bentall, 2003; Bick & Kinsbourne, 1987), is that the raw material of AVHs is internally generated inner speech. We go beyond these existing accounts, however, in claiming that inner speech is inherently dialogic. Two specific forms of dialogic inner speech are identified: Level 3 (expanded) inner speech, requiring high levels of VSM, and Level 4 (condensed) inner speech, associated with low-VSM demands. Specifically, it is Level 3 (expanded) inner speech that forms the raw material for AVHs. Such an account can account for the paradoxical 'alien-yet-self' quality of AVHs (Fernyhough, 2004), as well as fitting with the evidence from neural correlate studies that AVH-hearers differ from healthy controls only on tasks involving high levels of VSM.

The second challenge for any comprehensive account of AVHs is to explain why they are experienced as authored by a person other than one's self. Following Frith and Done (1988) and others, we attribute this feature of AVHs in part to a failure in a neurocognitive action self-monitoring system (Jones & Fernyhough, in press), which is put under particular pressure in high-VSM conditions, such as the generation of expanded inner speech. Beyond explaining errors of judgment about authorship, however, a comprehensive account of AVHs must meet a third challenge, namely to explain why voices are often perceived as having person-specific acoustic properties, distinct from those of the AVH-hearer. The Vygotskian approach to inner speech allows us to account for this third feature of AVHs with reference to the dialogic nature of inner speech. That is, human mentation involves the dialogic interplay of semiotically manifested perspectives on reality, such that our thinking is literally shot through with other voices. When combined with atypical VSM, of the kind that we and others propose to characterize AVH-hearers, in addition to further potential factors such as anterior cingulate and parietal cortex activation, these voices are perceived as having an alien quality.

The fourth challenge for a comprehensive theory is to explain why AVHs have characteristic contents and pragmatics, particularly their tendency to appear as second-person assertive utterances (e.g., commands). We have suggested that the developmental foundation of inner speech in private speech and external dialogue, where children gradually acquire linguistic means for regulating their own behavior, can account for the specific pragmatic qualities of AVHs.

We have set out a number of suggestions for how future studies involving AVH-hearers and healthy controls might test these hypotheses. We propose that further light may be shed on these issues through the use of a neglected population in AVH research. Many voice-hearers do not attract a clinical diagnosis such as schizophrenia and are able to deal with the experience of hearing voices in living normal lives (Romme & Escher, 1993). To our knowledge this population has not yet been invited to participate in neuroimaging studies. Our hypotheses would predict, for example, that neural correlates of expanded inner speech in healthy voice-hearers would be similar to those of AVH-hearers diagnosed with schizophrenia. In terms of existing neuroimaging paradigms, we would predict less SMA and temporal lobe activation than healthy controls when generating (particularly second-person) AVI. In addition, we note the need for much more research on the neural correlates of the experience of AVHs, to complement the rich body of research conducted with AVH-hearers in remission. There is also an urgent need to devise ecologically valid procedures for eliciting both forms of dialogic inner speech. Inner speech is, we suggest, a rich and heterogeneous phenomenon, and empirical studies of it are weakened by a continued reliance on unexamined assumptions about its form and nature.

References

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

Baddeley, A. D. (2001). Is working memory still working? *American Psychologist*, *56*, 851−864.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation*, *Vol. 8*. New York: Academic Press.

Balan, S., Spivak, B., Nachshoni, T., Kron, S., Mester, R., & Weizman, A. (1996). Auditory pseudohallucinations induced by a combination of hearing impairment and environmental stress. *Psychopathology*, *29*, 198−200.

Bentall, R. P. (1990). The illusion of reality: A review and integration of psychological research on hallucinations. *Psychological Bulletin*, *107*, 82−95.

Bentall, R. P. (2003). *Madness explained.* London: Penguin.

Berk, L. E. (1992). Children's private speech: An overview of theory and the status of research. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 17−53). Hove, UK: Lawrence Erlbaum Associates.

Bick, P. A., & Kinsbourne, M. (1987). Auditory hallucinations and subvocal speech in schizophrenic patients. *American Journal of Psychiatry*, *144*, 222−225.

Bullmore, E., Horwitz, B., Honey, G., Brammer, M., Williams, S., & Sharma, T. (2000). How good is good enough in path analysis of fMRI data? *NeuroImage*, *11*, 289−301.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215−222.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, *25*, 657−674.

Cooklin, R., Sturgeon, D., & Leff, J. (1983). The relationship between auditory hallucinations and spontaneous fluctuations of skin-conductance in schizophrenia. *British Journal of Psychiatry*, *142*, 47−52.

Creutzfeldt, O., Ojeman, G., & Lettich, E. (1989). Neuronal activity in the human lateral temporal lobe: II. Responses to the subject's own voice. *Experimental Brain Research*, *77*, 476−489.

David, A. S., & Lucas, P. A. (1993). Auditory verbal hallucinations and the phonological loop — A cognitive neuropsychological study. *British Journal of Clinical Psychology*, 431−441.

Dehaene, S., Artiges, E., Naccache, L., Martelli, C., Viard, A., Schurhoff, F., et al. (2003). Conscious and subliminal conflicts in normal subjects and patients with schizophrenia: The role of the anterior cingulate. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 13722−13727.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*, 970−974.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. D., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115*, 1753−1768.

Evans, C. L., McGuire, P. K., & David, A. S. (2000). Is auditory imagery defective in patients with auditory hallucinations? *Psychological Medicine*, *30*, 137−148.

Fernyhough, C. (1996). The dialogic mind: A dialogic approach to the higher mental functions. *New Ideas in Psychology*, *14*, 47−62.

Fernyhough, C. (2004). Alien voices and inner dialogue: Towards a developmental account of auditory verbal hallucinations. *New Ideas in Psychology*, *22*, 49−68.

Fernyhough, C. (2005). What is internalised? Dialogic cognitive representations and the mediated mind (Commentary on Tomasello et al.). *Behavioral and Brain Sciences*, *28*, 698−699.

Fernyhough, C. (submitted for publication). *Getting Vygotskian about theory of mind: Mediation, dialogue, and the development of social understanding*. Manuscript submitted for publication.

Fodor, J. (1975). *The language of thought.* Cambridge: Harvard University Press.

Ford, J. M., & Mathalon, D. H. (2005). Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations? *International Journal of Psychophysiology*, *58*, 179−189.

Ford, J. M., Mathalon, D. H., Kalba, S., Whitfield, S., Faustman, W. O., & Roth, W. T. (2001). Cortical responsiveness during inner speech in schizophrenia: An event-related potential study. *American Journal of Psychiatry*, *158*, 1914−1916.

Friedman, L., Kenny, J. T., Wise, A. L., Wu, D., Stuve, T. A., Miller, D. A., et al. (1998). Brain activation during silent word generation evaluated with functional MRI. *Brain and Language*, *64*, 231−256.

Frith, C. D., & Done, D. J. (1988). Towards a neuropsychology of schizophrenia. *British Journal of Psychiatry*, *153*, 437−443.

Frith, C. D., Rees, G., & Friston, K. (1998). Psychosis and the experience of self: Brain systems underlying self-monitoring. *Annals of the New York Academy of Sciences*, *15*, 170−178.

Haddock, G., Slade, P. D., Prasaad, R., & Bentall, R. P. (1996). Functioning of the phonological loop in auditory hallucinations. *Personality and Individual Differences*, *20*, 753−760.

Hermer-Vazquez, L., Spelke, E., & Katsnelson, A. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology*, *39*, 3−36.

Hoffman, R. E., & McGlashan, T. H. (1997). Synaptic elimination, neurodevelopment, and the mechanism of hallucinated "voices" in schizophrenia. *American Journal of Psychiatry*, *154*, 1683−1689.

Huang, J., Carr, T. H., & Cao, Y. (2001). Comparing cortical activations for silent and overt speech using event-related fMRI. *Human Brain Mapping*, *15*, 39−53.

Hubl, D., Koenig, T., Strik, W., Federspiel, A., Kreis, R., Boesch, C., et al. (2004). Pathways that make voices: White matter changes in auditory hallucinations. *Archives of General Psychiatry*, *61*, 658−668.

Hunter, M. D., & Spence, S. A. (2005). Left frontal activation. *British Journal of Psychiatry*, *187*, 89.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, *302*, 120−122.

Johns, L. C., Gregg, L., Allen, P., & McGuire, P. K. (2006). Impaired verbal self-monitoring in psychosis: Effects of state, trait and diagnosis. *Psychological Medicine*, *36*, 465−474.

Johns, L. C., Gregg, L., Vythelingum, N., & McGuire, P. K. (2003). Establishing the reliability of a verbal self-monitoring paradigm. *Psychopathology*, *36*, 299−303.

Jones, S. R. (in press). The neuropsychology of covert and overt speech: Implications for the study of private speech in children and adults. In A. Winsler, C. Fernyhough, & I. Montero (Eds.), *Private speech, executive functioning, and the development of verbal self-regulation*. Cambridge, UK: Cambridge University Press.

Jones, S. R., Fernyhough, C. (in press). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition*.

Lennox, B. R., Bert, S., Park, G., Jones, P. B., & Morris, P. G. (1999). Spatial and temporal mapping of neural activity associated with auditory hallucinations. *Lancet*, *353*, 644.

Lennox, B. R., Bert, S., Park, G., Medley, I., Morris, P. G., & Jones, P. B. (2000). The functional anatomy of auditory hallucinations in schizophrenia. *Psychiatry Research*, *100*, 13−20.

Leudar, I., & Thomas, P. (2000). *Voices of reason, voices of insanity: Studies of verbal hallucinations.* London: Routledge.

Leudar, I., Thomas, P., McNally, D., & Glinski, A. (1997). What voices can do with words: Pragmatics of verbal hallucinations. *Psychological Medicine*, *27*(4), 885−898.

Levine, D. N., Calvanio, R., & Popovics, A. (1982). Language in the absence of inner speech. *Neuropsychologia*, *20*, 391−409.

Luria, A. R. (1961). *The role of speech in the regulation of behavior.* Harmondsworth: Penguin.
Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*, 299−302.
Maudsley, H. (1886). *Natural causes and supernatural seemings.* London: Kegan Paul, Trench & Co.
McGuire, P. K., Shah, G. M. S., & Murray, R. M. (1993). Increased blood flow in Broca's area during auditory hallucinations in schizophrenia. *Lancet*, *342*, 703−706.
McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, *119*, 907−917.
McGuire, P. K., Silbersweig, D. A., Murray, R. M., David, A. S., Frackowiak, R. S. J., & Frith, C. D. (1996). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine*, *26*, 29−38.
McGuire, P. K., Silbersweig, D. A., Wright, I., Murray, R. M., David, A. S., Frackowiak, R. S., et al. (1995). Abnormal monitoring of inner speech: A physiological basis for auditory hallucinations. *Lancet*, *346*, 596−600.
Nayani, T. H., & David, A. S. (1996). The auditory hallucination: A phenomenological survey. *Psychological Medicine*, *26*, 177−189.
North, C. (1990). *Welcome, silence: My triumph over schizophrenia.* London: Arrow Books.
Ramirez, J. D. (1992). The functional differentiation of social and private speech: A dialogic approach. In R. M. Diaz & L. E. Berk (Eds.), *Private Speech: From Social Interaction to Self-regulation.* New Jersey: Lawrence Erlbaum Associates.
Reese, W. D. (1971). The hallucinations of widowhood. *British Medical Journal*, *210*, 37−41.
Romme, M., & Escher, S. (1993). *Accepting voices.* London: MIND Publications.
Seal, M. L., Aleman, A., & McGuire, P. K. (2004). Compelling imagery, unanticipated speech and deceptive memory: Neurocognitive models of auditory verbal hallucinations in schizophrenia. *Cognitive Neuropsychiatry*, *9*, 43−72.
Shergill, S. S., Brammer, M. J., Amaro, E., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2004). Temporal course of auditory hallucinations. *British Journal of Psychiatry*, *185*, 516−517.
Shergill, S. S., Brammer, M. J., Fukuda, R., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2003). Engagement of brain areas implicated in processing inner speech in people with auditory hallucinations. *British Journal of Psychiatry*, *182*, 525−531.
Shergill, S. S., Brammer, M. J., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2000). Mapping auditory verbal hallucinations in schizophrenia using functional magnetic resonance imaging. *Archives of General Psychiatry*, *57*, 1033−1038.
Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine*, *31*, 241−253.
Shergill, S. S., Bullmore, E., Simmons, A., Murray, R., & McGuire, P. (2000). Functional anatomy of auditory verbal imagery in schizophrenic patients with auditory hallucinations. *American Journal of Psychiatry*, *157*, 1691−1693.
Silbersweig, D. A., Stern, E., Frith, C., Cahill, C., Holmes, A., Grootoonk, S., et al. (1995). A functional neuroanatomy of hallucinations in schizophrenia. *Nature*, *378*, 176−179.
Soltysik, D. A., & Hyde, J. S. (2006). Strategies for block-design fMRI experiments during task-related motion of structures of the oral cavity. *NeuroImage*, *29*, 1260−1271.
Spence, S. A., Brooks, D. J., Hirsch, S. R., Liddle, P. F., Meehan, J., & Grasby, P. M. (1997). A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain*, *120*, 1997−2011.
Stephane, M., Barton, S., & Boutros, N. N. (2001). Auditory verbal hallucinations and dysfunctions of the neural substrates of speech. *Schizophrenia Research*, *50*, 61−78.
Vygotsky, L. S. (1978). In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher mental processes.* Cambridge, MA: Harvard University Press. Original work published 1930, 1933, and 1935.
Vygotsky, L. S. (1987). *Thinking and speech. The Collected Works of L.S. Vygotsky*, *Vol. 1*. New York: Plenum. Original work published 1934.
Vygotsky, L. S. (1997). Genesis of higher mental functions. In R. W. Rieber (Ed.), *The Collected Works of L.S. Vygotsky*, *Vol. 4*. New York: Plenum. Original work published 1931.
Watson, J. B. (1920). Is thinking merely the action of language mechanisms? *British Journal of Psychology*, *11*, 87−104.
Weiskrantz, L. (1988). *Thought without language.* Oxford: Oxford University Press.
Wertsch, J. V. (1980). The significance of dialogue in Vygotsky's account of social, egocentric and inner speech. *Contemporary Educational Psychology*, *5*, 150−162.
Winsler, A. (2004, November). *Still talking to ourselves after all these years: Vygotsky, private speech, and self-regulation*. Invited address given at First International Symposium on Self-Regulatory Functions of Language, Madrid, Spain.
Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., & Frackowiak, R. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain*, *119*, 1803−1817.
Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*, 846−849.
Zatorre, R. J., & Halpern, A. R. (1993). Effect of unilateral temporal lobe exclusion on perception and imagery of songs. *Neuropsychologia*, *31*, 221−232.