Hearing Voices

The Histories, Causes and Meanings of Auditory Verbal Hallucinations

Simon McCarthy-Jones

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by
Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9781107007222

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data

McCarthy-Jones, Simon, 1978–
Hearing voices : the histories, causes, and meanings of auditory verbal hallucinations / Simon McCarthy-Jones.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-107-00722-2 (Hardback)
I. Title.
[DNLM: 1. Hallucinations–history. 2. Hallucinations–psychology. 3. Auditory Perception–physiology. 4. Hallucinations–etiology. 5. Mental Disorders–history. 6. Public Opinion–history. WM 204]
616.89–dc23

2011035576

ISBN 978-1-107-00722-2 Hardback

For those whose voices have not been heard (yet)
With gratitude to my wife, whose love creates me

Introduction to Part III

Having established in the previous chapters what the phenomenology of AVHs is like, we are now in a position to try and attempt to explain what causes these experiences. For any who question the need to understand the causes of voices, we may refer to the words given to the Merovingian by the Wachowski brothers in the film *The Matrix Reloaded*. '*Why* is the only source of real power, without it you are powerless' (Wachowski & Wachowski, 2003).

Chapter 8 will examine the neural underpinnings and biology of AVHs.

Chapters 9 and 10 will then examine the cognitive psychology of AVHs, being guided by both the phenomenology of AVHs as well as the findings from Chapter 8. Chapter 9 will focus specifically on inner speech models, with Chapter 10 looking predominantly at memory and hypervigilance models.

Chapter 11 will go beyond the confines of an individual's head to examine environmental factors that may cause someone to hear voices, and how these work through an individual's biology.

8 Neuroscience and hearing voices: it's the brain, stupid?

Before we begin this chapter, a word on terminology. First, I am going to use the terms 'voice-hearing brain', SZ:AVH+ (an individual diagnosed with schizophrenia and who hears voices) and SZ:AVH− (an individual diagnosed with schizophrenia but who does not hear voices). Such phrases have the potential to objectify, and similar phrases, such as Bartels-Velthuis, Jenner & van de Willige's (2010) term 'AVH-positive children' (p. 43) suggest the voice-hearer has tested positive for a disease. In regard to the term 'voice-hearing brain', it will quite rightly be pointed out that brains do not hear voices, people do. Also, given that in Chapter 7 we saw AVHs occur in a number of people in the general population, all brains could be seen as potentially voice-hearing brains under the right (or wrong) conditions. However, the terminology I employ here is simply meant to be a convenient shorthand way of referring to participants in the studies, and to make the text less cumbersome. It should also be noted that use of this terminology is not to endorse the concept of schizophrenia, whose reliability and validity has been questioned (Bentall, 2003; Boyle, 2002), although the distress and suffering of individuals with such a diagnosis is very real.

Getting our bearings

Before we start on this journey, it may help to get our bearings in the brain. Figure 8.1 shows a side-on diagram of the cortical regions of the brain (the front on the left, the back on the right), and, importantly for the purposes of models of AVHs, indicates Broca's area (typically associated with overt speech and inner speech production), Wernicke's area (typically associated with speech perception) and the arcuate fasciculus white matter pathway that links them together. Although there is also another indirect pathway that links Broca's and Wernicke's area, which runs through the inferior parietal cortex (Catani, Jones & ffytche, 2005), this pathway has been much less studied in relation to AVHs, and so will not be focused upon here. In Figure 8.1 the four lobes of the

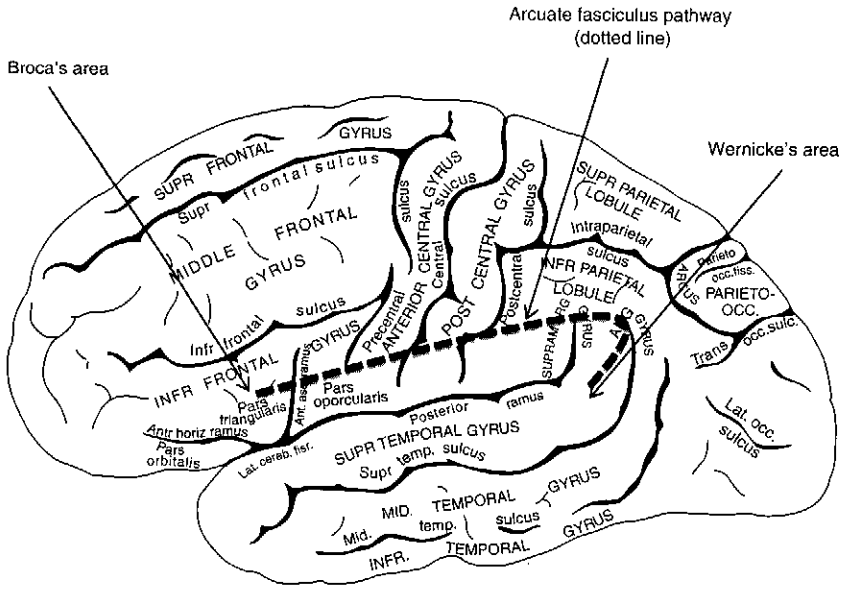

Figure 8.1. The cortex of the brain
Source: Taken from Wikipedia Commons at http://commons.wikimedia.org/wiki/File:Mediaal_(gray).PNG

brain can also be broadly determined, being the frontal lobe (broadly corresponding to the areas labelled superior, middle and inferior frontal gyri), the temporal lobe (broadly corresponding to the areas labelled superior, middle and inferior temporal gyri), the parietal lobe (the superior and inferior parietal lobules) and the occipital lobe (broadly located around the area marked 'lat occ. sulcus').

Early neuro-stimulation studies

A relatively simple way to find out what brain areas might be involved in AVHs is to electrically simulate the brain directly and see whether it causes AVHs. This is precisely what Penfield & Perot (1963) did. During neurosurgery on 520 patients with epilepsy, which involved removing portions of the scalp in order to be able directly to access the brain, they were able to use an electrode to electrically stimulate the brain directly. As the patients had only received local anaesthetics to their scalps and the brain itself does not feel pain, the patients were wide awake and able to report on what they saw, felt or heard during this process.

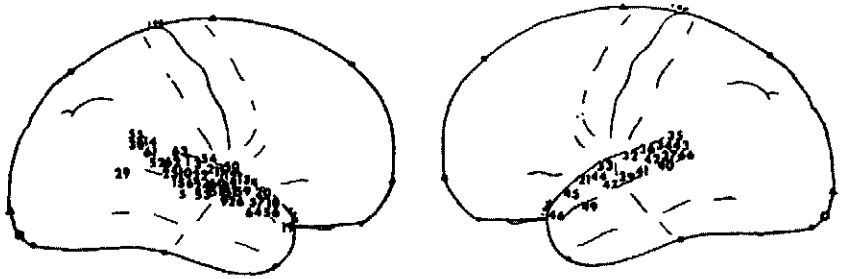

Figure 8.2. Areas of direct brain stimulation resulting in auditory hallucinations
Source: Reproduced from Penfield, W. & Perot, P. (1963). The brain's record of auditory and visual experience. *Brain*, 86, 595–696, by permission of Oxford University Press.

Before their study, in a review of the case histories of 1,132 of their patients suffering from focal epilepsy, Penfield & Perot had identified no patients with auditory hallucinations who had their focus of epileptic discharge outside of the temporal lobe. This suggested to them that stimulating the temporal lobe would be likely to produce auditory hallucinations. Consistent with their prediction, they found that only stimulation of a specific temporal lobe structure, the superior temporal gyrus (STG), resulted in auditory hallucinations. Overall, Penfield & Perot elicited from the STG either a voice, voices, music or a meaningful sound from 66 different stimulation points in 24 patients. Of these auditory experiences, they note: 'a voice or voices was the most common response' which resulted from stimulation of 46 cortical points. Of these 46 points, 31 resulted in voices that were recognised/known to the patient. There was a trend for voices to result more often from STG stimulation of the non-dominant hemisphere (i.e. in right-handed people, the right hemisphere). Figure 8.2 shows the specific areas in the right and left hemisphere that, when stimulated, gave rise to auditory hallucinations.

Within the STG, if the primary auditory cortex (and Heschl's gyrus) was stimulated, patients reported hearing non-linguistic simple sounds, such as buzzing or whistling. However, they report in some cases that when the electrode was then moved towards Wernicke's area, actual voices were heard. For example, they report that 'In Case 29 ... an electrode was inserted into the anterior transverse temporal gyrus and the patient reported, "a buzzing". As the electrode was withdrawn, so that it came into the cortex of the first temporal gyrus, he exclaimed, "Someone is calling"' (p. 666). Stimulation of the STG led to a range of AVHs, which included patients reporting 'I heard voices' (p. 651),

'my mother and father talking' (p. 645), and 'something like a crowd' (p. 640). In addition to AVHs, musical hallucinations and forms of non-verbal auditory hallucinations (e.g. bangs) were also caused. Notably, only a small number of the total number of stimulations of the temporal lobe (7.7%) actually led to any form of hallucination.

Mahl *et al.* (1964) also investigated the effects of direct stimulation of the left temporal lobe, although they only used a single patient with epilepsy. On one occasion the patient reported hearing a conversation between two people. At another time she reported hearing some kind of offensive word. Although I have used the word 'hear', interestingly the authors report that when questioned in detail, 'the patient cannot decide whether she actually heard people saying words to her or whether she thought about the words' (p. 347). Furthermore, the authors note that what she heard were 'not merely words which she associates with past events; the words, sentences, and conversations occur at the time she reports them' (ibid.). Yet, the authors conclude that 'we cannot tell whether the hallucinatory experiences were the equivalent of dreams or psychotic hallucinations or the equivalent of an exact playback' (p. 357). Ferguson *et al.* (1969) also reported upon the effects of direct electrical stimulation of the brain, but in this case specifically the left amygdala region and anterior hippocampus. This did not result in any AVHs, although visual hallucinations did occur. Consistent with the above pattern of findings, the use of the more recent and advanced magnetoencephalography (MEG: which examines the magnetic field change associated with changes in electrical activity in the brain) technique has found that AVHs occurring in epilepsy are associated with spikes of activity in the STG (Mohamed *et al.*, 2006). Finally, it is also worth noting an unexpected case of AVHs caused by stimulation of an area long thought to have an involvement in AVHs, but not reported on by Penfield & Perot (1963) as causing AVHs when directly stimulated. Lesser *et al.* (1984) report that direct electrical stimulation of Broca's area caused a patient to report AVHs, which spoke in the patient's voice and said single words, phrases or sentences. Nevertheless, the authors were careful to note that stimulation of this area may have led to activation in temporal lobe regions, which in turn could have been the cause of the patient's AVH.

What can we conclude from such studies as to the causes of AVHs? Is it likely that spontaneous epileptic electrical activation in the superior temporal gyrus is a model for all AVHs? First we may note that David (1994), after reviewing the phenomenology of the AVHs Penfield & Perot (1963) elicited, notes that 'it is striking how *unlike* the reports [of AVHs during electrical stimulation] are of auditory hallucinations described by psychotic patients' (p. 271, emphasis added). Specifically, David argues that only 5 of the 40 cases reported by Penfield & Perot show 'verbatim accounts of clear [AVHs]' (ibid.). However, this does not necessarily invalidate the model, as not all AVHs in patients diagnosed with psychosis are clear (Chapter 4), and we may note two other cases in addition to the ones David reports: case 15, 'voices yelling at me' (p. 630), and case 29, 'a man's voice, I could not understand what he said' (p. 640). As another argument against the phenomenological similarity between these induced AVHs and those found in patients diagnosed with psychosis, David also observes that two-thirds of patients in the Penfield & Perot study were able to identify their AVHs as voices from their past (indeed, Penfield & Perot themselves note that 'the patient usually recognizes it [the induced AVHs] spontaneously as coming from his past'). We may further add that we do not see many phenomenological parallels in the AVHs created in these studies to common properties of the voices of individuals diagnosed with psychosis (or healthy voice-hearers), such as commands. However, some common properties can be seen (e.g. being shouted at, and voices directly addressing the patient). Yet overall, it appears that spontaneous focal epileptic activity in the STG is likely not to be a good model of all AVHs. However, this may be a good model for the Type 2 (Static) AVHs highlighted in Chapter 7.

Before proceeding, it is worth considering if we can even model epileptic AVHs as resulting simply from focal epileptic activity in the STG. We saw in Chapter 4 that AVHs in non-psychotic patients with epilepsy have a phenomenology different from that of individuals diagnosed with psychosis in many ways (e.g. commands are rare) (Korsnes *et al.*, 2010). But even though such AVHs appear to be phenomenologically distinct from those found in psychosis, on a dichotic listening task non-psychotic epilepsy patients with inter-ictal AVHs have been found to have patterns of neural activation similar to that of SZ:AVH+ (ibid.). This suggests that AVHs in epilepsy may involve more than simply activations in just the STG. More research thus remains to be done on commonalities and differences in both the phenomenology of AVHs and the neural activation during AVHs, between patients with epilepsy (without psychosis) and those found to be typical of a cluster of psychiatric diagnoses (and in people without psychiatric diagnoses) in Chapter 4.

The neuroscience of hearing voices

Although we saw in Chapter 4 that the phenomenology of AVHs in patients diagnosed with psychosis was much like that of individuals with

AVHs who had been given other psychiatric disorders, or indeed the AVHs of healthy voice-hearers, the rest of this chapter, after a few brief comments, will focus solely on studies employing patients diagnosed with schizophrenia with AVHs (SZ:AVH+). I do this in order to allow comparability across studies, as inclusion of studies of AVHs in individuals diagnosed with other psychiatric disorders, due to the different medication regimes of such patients, may cause confounds and confuse the emerging picture of neural activation/structural changes uniquely associated with AVHs. However, it is likely that the neural processes involved in SZ:AVH+ are the same as those in individuals with phenomenologically similar AVHs but with different (or no) psychiatric diagnoses. For example, a study published just as I was making the final changes to this book (and hence not included in the analyses in the rest of this chapter) found that the neural regions activated during AVHs in SZ:AVH+ were not different to that during AVHs in healthy voice-hearers (Diederen *et al.*, in press). Similarly, a study by Knöchel *et al.* (in press), discussed in more detail later, found that changes to the corpus callosum were associated with hallucination severity (although AVHs specifically were not reported on) in patients diagnosed with schizophrenia, and that a trend towards a similar pattern was also found in relation to hallucination-proneness in healthy individuals. Furthermore, (a) AVHs in patients with epilepsy diagnosed with psychosis have been found to be associated with a wide range of neural changes (for example, fronto-temporal white matter abnormalities, Flugel *et al.*, 2006), (b) AVHs in Parkinson's disease have been found to be associated with hypoperfusion in the bilateral prefrontal cortex and right STG (Matsui *et al.*, 2006), and (c) AVHs in patients with Alzheimer's disease has been found to be associated with smaller parahippocampal gyrus volumes (Forstl *et al.*, 1994), all areas that will be highlighted as important in the review below of AVHs in SZ:AVH+. Although more work does remain to be done both on phenomenological and neural differences in AVHs between these psychiatric diagnoses (and to link specific aspects of phenomenology of AVHs to neural activations), it appears that such findings at least support a symptom-based approach to AVHs.

Structural neuroimaging studies of AVHs

Structural imaging studies of the brain using magnetic resonance imaging can shed light on how the brain of voice-hearers is different to that of non-voice-hearers. Whilst older studies focused on grey matter structures in the brain (which contain the nerve cell bodies themselves), more recent studies have been able to examine white matter tracts in the brain (responsible for the transmission of signals between grey matter areas, and which may both cause activity in grey matter structures to be either excited or inhibited).

Table 8.1. *Structural studies comparing SZ:AVH+ to healthy controls*

Study	N (SZ/HC)	Method	Superior temporal gyrus	Insula
Hubl *et al.* (2010)	13:13	MRI (DTI)	↑ (right)	–
Cachia *et al.* (2008)	30–28	MRI (VBM)	↓ (bilateral)*	–
Garcia-Martí *et al.* (2008)	18–11	MRI (VBM)	↓ (bilaterally)	↓ (bilaterally)
Hugdahl *et al.* (2008)	6–12	MRI (VBM)	↓ (left)	X
O'Daly *et al.* (2007)	28–32	MRI (VBM)	↓ (right)	↓ (bilaterally)
Shapleske *et al.* (2002)	41–32	MRI (VBM)	X	↓ (bilaterally)
Shapleske *et al.* (2001)	74–32	MRI (ROI)	X	–
Havermans *et al.* (1999)	15:17	MRI (ROI)	X	–

* Sulcation in this region was lower.
– = area not examined, SZ = patients diagnosed with schizophrenia, HC = healthy controls, X = no differences found.
Source: Compiled by the author.

Structural imaging studies of grey matter in SZ:AVH+

To date there have been 15 structural imaging studies which have examined grey matter (GM) volumes in relation to AVHs in patients diagnosed with schizophrenia. These are grouped below based on the nature of the analyses they performed.

SZ:AVH+ compared to healthy controls (HCs)

Eight studies have employed this design (Table 8.1). The only replicated findings are that SZ:AVH+ have reduced STG GM volumes (found in 4 of 8 studies) and reduced insula volumes (3 of 4 studies) compared to HCs. Yet the most recent study (Hubl *et al.*, 2010) found *increased* STG volumes in SZ:AVH+ compared to HCs. An obvious limitation of comparing SZ:AVH+ to HCs is that structural changes identified may not be specific to the presence of AVHs but to factors associated with a diagnosis of schizophrenia, or may result from the documented effects of antipsychotic medication on both cortical and sub-cortical GM volumes (Dazzan *et al.*, 2005; Gur *et al.*, 1998). For example, Vernon *et al.* (2010) found, in rats, that antipsychotic administration (both haloperidol and

Table 8.2. *Grey matter volumes in participants with SZ:AVH+ compared to SZ:AVH−*

Study	N (SZ:HC)	Method	Superior temporal gyrus
Hubl *et al.* (2010)	13:13	MRI (DTI)	↑ (right)
Shin *et al.* (2005)	17:8	MRI (VBM)	–
Onitsuka *et al.* (2004)	13:6	MRI (ROI)	↓ (left)
Shapleske *et al.* (2002)	41:31	MRI (VBM)	X
Shapleske *et al.* (2001)	44:30	MRI (ROI)	X
Havermans *et al.* (1999)	15:15	MRI (ROI)	X
DeLisi *et al.* (1994)	Not reported	MRI (ROI)	X

olanzapine) caused a 6–8 per cent decrease in whole brain volume, driven mainly by a reduction in frontal cerebral cortex volume. We thus need to turn to studies which compare patients diagnosed with schizophrenia with and without AVHs, which will be more useful, as both will be on antipsychotic medication, eliminating this as a potential confound.

SZ:AVH+ vs SZ:AVH−

Seven studies have been done of this design (Table 8.2) which attempt to identify structural changes specific to AVHs, rather than schizophrenia per se. The only area identified by more than one study has been the STG. However, whereas Onitsuka and colleagues (2004) found smaller STG volumes specific to SZ:AVH+, this was not replicated by the six other studies, five of which found no differences, and one which found an increased volume specifically in Heschl's gyrus. Inconsistencies in these studies may result from the small number, the relatively small sample sizes often employed, the potentially confounding presence of varying levels of hallucinations in other modalities/other positive symptoms and inconsistencies in the definitions of SZ:AVH+ and SZ:AVH− used. For example, whereas one study (Shapleske *et al.*, 2001) defined SZ:AVH− as those with a score less than 2 on the Scale for Assessment of Positive Symptoms (SAPS) (Andreasen, 1984) Auditory Hallucination item, another (Onitsuka *et al.*, 2004) defined SZ:AVH− as those with a zero score on this measure.

AVH severity correlated with GM volumes

Eight studies of this design have been performed (Table 8.3). Studies have used a range of measures of AVH severity, including the

Table 8.3. *Correlations of severity of AVHs with grey matter volumes*

Study	N	Method	Superior temporal gyrus	Inferior frontal gyrus	Postcentral gyrus
Nenadic *et al.* (2010)	99	MRI (VBM)	↓ (bilaterally)	X	↓ (left)
Modinos *et al.* (2009)	26	MRI (VBM)	X	↑ (left)	X
García-Martí *et al.* (2008)	18	MRI (VBM)	X	↓ (left)	↓ (right)
Gaser *et al.* (2004)	85	MRI (DBM)	↓	↓ (right)	
Shapleske *et al.* (2001)	74	MRI (VBM)	X	–	–
Levitan *et al.* (1999)	30	MRI (ROI)	↓ (left)	–	–
Cullberg *et al.* (1992)	33	CT	–	–	–
Barta *et al.* (1990)	15	MRI (ROI)	↓ (left)	–	–

PSYRATS-AH (Haddock *et al.*, 1999), the SAPS Auditory Hallucination item and the sum of the SAPS Auditory Hallucination, Voices Commenting and Voices Conversing items. The most consistent finding was that four of the seven studies examining the STG found smaller STG volumes to be associated with more severe AVHs. Of the four studies examining the inferior frontal region, two found reduced volumes to be associated with more severe AVHs, one found an increased volume to be associated with more severe AVHs, and one found no relation. The only other replicated finding was that two studies (out of three) found reductions in the post-central gyrus GM volume to be associated with increased levels of AVHs. One further study (Cachia *et al.*, 2008) examined the correlation between AVH severity and measures of gyrification and sulcation in the temporal lobe, but failed to find any significant correlations.

Structural imaging studies of white matter and connectivity in AVHs

One initial source of evidence that white matter (WM) problems are related to AVH comes from studies of the condition metachromatic leukodystrophy, a demyelinating genetic disorder which occurs in about 1/40,000 people (Black, Taber & Hurley, 2003). The loss of WM is most prominent in the frontal and parietal lobes, and this is associated with hearing voices and a range of other experiences (ibid.).

The WM tract most likely to be of relevance to AVHs (and hence the most studied) is the arcuate fasciculus (AF). The AF is the main part of the larger superior longitudinal fasciculus (SLF) which connects the front and back of the brain (Bernal & Ardila, 2009). The lateral part of

the AF contains shorter U-shaped fibres that connect a range of areas, including Broca's area to Wernicke's area (Hubl *et al.*, 2004) via a possible relay station in the premotor or motor cortex (Bernal & Ardila, 2009).

Four studies have simply examined WM volume (rather than tract integrity) in those with AVHs. Compared to HCs, O'Daly *et al.* (2007) found that SZ:AVH+ had reduced WM volume in the left but not the right STG, as well as in the right inferior longitudinal fasciculus. Although Shapleske *et al.* (2002) when comparing SZ:AVH+ to HCs also found reduced WM in the right frontal lobe near the SLF, and increased WM in the left temporal-parietal connecting tracts, when they compared SZ:AVH+ to SZ:AVH− they found no differences. Similarly, Rossell *et al.* (2001) found no difference in the size of the CC between SZ:AVH+ and SZ:AVH−. However, Seok *et al.* (2007) did find a WM volume difference specific to SZ:AVH+, with WM density in the left SLF being significantly increased in SZ:AVH+, but not different in SZ:AVH− compared to HCs.

Going beyond simple volumetric measurements of WM, diffusion tensor imaging (DTI) allows the assessment of the integrity of WM tracts in the brain. When water is within WM tracts, its direction of diffusion is primarily limited to being parallel to it (Basser, Mattiello & Lebihan, 1994). The degree of restriction on the direction of flow of water molecules is termed fractional anisotropy (FA). DTI works by assessing the FA of water in WM tracts. Lower relative FA indicates damaged WM (Le Bihan *et al.*, 2001), including loss of structural organisation and the expansion of extracellular space (Dong *et al.*, 2004). Increased relative FA, however, can be taken to mean increased anatomical connectivity (Hubl *et al.*, 2004).[1]

Five DTI studies have focused specifically on AVHs. As shown in Table 8.4, Hubl and colleagues (2004) and Rotarska-Jagiela and colleagues (2009) both found increased FA values in the arcuate fasciculus in SZ-AVH+ compared to HCs. Yet de Weijer *et al.* (2011), Seok *et al.* (2007) and Shergill *et al.* (2007) found lower FA values in SZ:AVH+ compared to HCs (note, the Shergill study is not cited in the SZ:AVH+ vs HCs section of Table 8.4, as only two-thirds of patients, not all, had AVHs). The only studies (Hubl *et al.*, 2004; Seok *et al.*, 2007) to have compared SZ:AVH+ with SZ:AVH− found increased FA in the SLF in

[1] Basically, imagine water flowing in a straight pipe. If you make holes in the side of the pipe and damage it, the water will start spurting out in all directions (low relative FA), whereas if the pipe has no holes in the side and is nice and smooth inside, the water will all flow in the same direction along the pipe (high relative FA).

Table 8.4. *Diffusion tensor imaging studies of AVHs*

Study	N	Arcuate fasciculus	Cingulum	Corpus callosum
FA in SZ:AVH+ compared to HCs				
de Weijer *et al.* (2011)	44SZ-AVH+ 42HC	↓ (bilaterally)	X	–
Rotarska-Jagiela *et al.* (2009)	12SZ-AVH+ 12HC	↑ (bilaterally)	X	↓
Seok *et al.* (2007)	15SZ-AVH+ 22HC	↓ (left)	↓	X
Hubl *et al.* (2004)	13SZ-AVH+ 13HC	↑ (temporoparietal region, bilaterally)	X	–
		↓ (all other parts)		↓
FA in SZ:AVH+ vs SZ:AVH−				
Seok *et al.* (2007)	15SZ-AVH+ 15SZ-AVH−	↑ (left frontal portion)	↑ (left caudal)	X
Huble *et al.* (2004)	13SZ-AVH+ 13SZ-AVH−	↑ (all parts bilaterally)	↑	↑
Correlation between AVH severity and FA values				
Rotarska-Jagiela *et al.* (2009)	12SZ-AVH+	+ve (bilateral)	−ve (left)	X
Seok *et al.* (2007)	15SZ-AVH+	+ve (left frontal)	+ve (anterior)	X
Shergill *et al.* (2007)	33SZ	+ve (bilateral)	+ve (anterior)	X

Source: Compiled by the author.

SZ:AVH+. In addition to the DTI studies presented in Table 8.4, a further study examined functional connectivity between neural areas, comparing SZ:AVH+ and HCs, and assessing concurrent activation in these areas. This study by Vercammen *et al.* (2010a) found significantly decreased connectivity between the left temporoparietal junction (TPJ) and the right homotope of Broca's area in SZ:AVH+ compared with HCs.

A number of studies have examined the relation between WM connectivity and AVH severity. All studies employing DTI and examining the AF found that AVH severity was associated with increased FA (Table 8.4). The majority of studies also found a positive correlation between AVH severity and FA in the cingulum. Although all studies found that AVH severity in SZ:AVH+ correlated with increased FA values in the AF, when compared to HCs, the studies were divided as to whether SZ:AVH+ had higher or lower absolute levels of FA in this region. It is somewhat paradoxical that, for example, Seok and colleagues (2007)

found that SZ:AVH+ had lower FA values in the AF than HCs, but within the SZ:AVH+ group higher FA values (i.e. those nearer the scores of HCs) had more severe AVHs. On balance, the following picture emerges. It appears that, compared to healthy controls, SZ:AVH+ have lower FA in the AF (here I give weight to the largest study in this area, coming from the excellent Dutch research group of Iris Sommer and colleagues), but that this is associated with a diagnosis of schizophrenia per se. Within the group of patients diagnosed with schizophrenia, higher FA in the AF is associated with more severe voices. Thus patients diagnosed with schizophrenia have an absolute hypo-connection in the AF, but hearing voices within this group is associated with relative hyper-connection in the AF. Confused? Me too.

Although the studies reported in Table 8.4 did not reliably find FA in the CC to be specifically related to AVHs, a recent study by Knöchel *et al.* (in press) has found that lower CC volumes and fibre integrity in patients diagnosed with schizophrenia are associated with more severe hallucinations (the scale they used was a measure of hallucinations per se, rather than specifically AVHs, and hence this study is not included in Table 8.4). Finally, it is notable that connectivity problems may be related to specific types of voices. For example, Lee *et al.* (2009) found FA values in the left STG correlated positively with levels of voices commenting (but not conversing). This again points us to the potential existence of sub-types of voices with different underlying neural mechanisms as proposed in Chapter 7. Voices commenting may fit into my class of Type 1b voices, underpinned by neural connectivity problems with inner speech (Chapter 9), and voices conversing may fit into my class of Type 2 voices (specifically Type 2b voices – see Chapter 11), underpinned by spontaneous, epileptic-like activity in the STG. This, of course, remains speculative.

In addition to these DTI studies, a number of studies have examined functional connectivity between neural areas in voice-hearers. Vercammen *et al.* (2010a) studied the functional connectivity between neural areas in SZ:AVH+ while at rest. It was found that AVH severity correlated negatively with functional connectivity between the left TPJ and the bilateral anterior cingulate, as well as with functional connectivity between the left TPJ and the bilateral amygdala. Gavrilescu *et al.* (2010) found that, when at rest, inter-hemispheric connectivity between both left and right primary auditory cortex, and left and right secondary auditory cortex, was worse in SZ:AVH+ than both HCs and SZ:AVH−. In addition, Hashimoto *et al.* (2010) found that the severity of AVHs was correlated with the functional connectivity between the left sylvian-parietal-temporal area and the anterior insula, a pathway

which the authors note is critical for speech production (Hickok & Poeppel, 2007). Moving away from studies while voice-hearers are 'at rest', Raij *et al.* (2009) failed to find any regions of the brain that coupled (i.e. covaried) with Broca's area differently during AVHs and during non-AVH periods in the scanner. However, they did find that during AVHs, coupling between left Broca's area and the bilaternal supratemporal cortex, right posterior temporal lobe and middle right anterior cingulate cortex, all correlated with the subjective reality of patients' voices. Furthermore, during AVHs the coupling between Broca's area and the right posterior STG, as well as Heschl's gyrus in the left auditory cortex, also correlated with the perceived reality of voices. The reality of AVHs correlated negatively with the coupling between Broca's area and the posterior cingulate cortex. As we will see later, such findings are consistent with models in which (inner) speech production areas fail to communicate with other regions of the brain (such as speech perception areas), resulting in AVHs.

Functional neuroimaging studies of AVHs

Activity during AVHs

Sixteen studies (4 PET, 12 fMRI) have examined the neural activity in SZ:AVH+ associated with the presence of AVHs, compared to non-AVH periods (Table 8.5). Overall, 69 per cent of studies found activation in the STG, 63 per cent in cingulate regions, 56 per cent in the middle temporal gyrus, 50 per cent in the inferior frontal gyrus (IFG), and 25 per cent in the parahippocampal gyrus, insula, middle frontal gyrus and cerebellum. Activation in these regions was not consistently found to be specific to a given hemisphere, or to occur bilaterally, though. A limitation of all but two of the studies was their sample size. Today it is generally accepted that a sample size of more than twenty participants is required for a functional magnetic resonance imaging (fMRI) study to produce reliably replicable findings (Thirion *et al.*, 2007). However, only two studies had a sample size in excess of this (Diederen *et al.*, 2010; Sommer *et al.*, 2008). Areas identified by both of these studies were bilateral activation in the IFG (primarily in the right hemisphere), the STG, insula, supramarginal gyrus and cerebellum.

Activity preceding AVHs

As studies of neural activation during AVHs cannot distinguish between activity resulting in their production and activity resulting from their

Table 8.5. *Functional imaging studies of AVHs in SZ:AVH+*

Study, imaging method (N)	AH assessment	IFG	MFG	Insula	STG	MTG	Cingulate	Parahippo-campal cortex	Other regions
Diederen *et al.* 2010, fMRI (24)	Balloon squeeze	Bilateral	Right	Bilateral	Bilateral	Bilateral			Post- and pre-central gyrus, inferior parietal lobule, cerebellum , bilateral supermarginal and superior frontal gyrus
Raij *et al.* 2009, fMRI (11)	Button press	Bilateral					Right	Right	Right posterior and left anterior temporal lobe
Sommer *et al.* 2008, fMRI (24)	Balloon squeeze	Bilateral		Bilateral	Bilateral		Left		Medial frontal and supramarginal gyrus, right cerebellum
Copolov *et al.* 2003, fMRI (8)	Button press	Right			Left	Right	Left posterior	Left	Right medial frontal
Shergill *et al.* 2001, fMRI(1)	Post-scan report			Right	Right	Right		Right	
Shergill *et al.* 2000, fMRI (6)	Post-scan report	Left	Bilateral	Bilateral		Bilateral	Anterior bilateral	Left	Right central gyrus, thalamus, parietal lobule, left hippocampus
Van der Ven *et al.* 2005, fMRI (6)	Button press				Bilateral HG*				
Lennox *et al.* 2000, fMRI (4)	Button press		Left		Bilateral	Bilateral			
Dierks *et al.* 1999, fMRI (3)	Button press	Bilateral			Left, incl HG	Left			Left amygdala, hippocampus
Lennox *et al.* 1999, fMRI (3)	Button press	Right	Right		Bilateral	Right	Right anterior		Right cuneus
Bentaleb *et al.* 2002, fMRI (1)	AVH prevented by noise				Left PAC	Right			
Woodruff *et al.* 1995, fMRI (1)	Finger raise				Right	Right	Right anterior		Thalamus. Right prefrontal, premotor and parietal cortex
McGuire *et al.* 1993, PET (12)	Symptomatic v remitted	Left					Left anterior (trend level)		
Parellada *et al.* 2008, PET (9)	Symptomatic v remitted						Anterior		SMA, medial superior frontal area, cerebellum
Silbersweig *et al.* 1995, PET (6)	Button press						Right anterior	Bilateral	Bilateral thalamus, right putamen and caudate, left cerebellum
Suzuki *et al.* 1993, PET (5)	Symptomatic v remitted				Left		Anterior		

* In 50% of patients.
HG = Heschl's gyrus, PAC = primary auditory cortex

consequences (Hoffman *et al.*, 2008a), studies have also examined the activity in the seconds immediately preceding onset of AVHs. In the largest of the four studies done in this area, Diederen *et al.* (2010: N = 15) examined neural activation in the six-seconds preceding AVHs, the onset of which SZ-AVH+ indicated by squeezing a balloon, compared to the six second periods preceding random balloon squeezes (not followed by AVHs). AVHs were most prominently preceded by deactivation in the left parahippocampal gyrus, and also by deactivation in the left STG, right inferior frontal (Broca's area) and left middle frontal gyri, right insula and left cerebellum. The finding of deactivation in the parahippocampal gyrus was also found in an earlier, smaller study by Hoffman *et al.* (2008a: N = 6) in the period 4.5–1.5s before AVHs. However, Hoffman and colleagues also found deactivation in the anterior cingulate, and activation, not deactivation, of the insula. Although Hoffman and colleagues suggested that insula activation may be due to preparation for motor movement to signal AVH-onset, Diederen and colleagues' study, which also found insula involvement whilst controlling for motor movements, suggests this is not the case. Diederen and colleagues' finding of a role for the inferior frontal gyrus was not replicated by Hoffman and colleagues, however Shergill and colleagues (2004: N = 2) did report activation of the inferior frontal gyrus (although the left, and not the right) preceding AVHs.

A consistent finding of earlier smaller studies (Hoffman *et al.*, 2008a; Lennox *et al.*, 1999; Shergill *et al.*, 2004), not found by Diederen and colleagues (2010), was activation around the middle temporal region preceding onset of AVHs specifically the right middle temporal region preceding onset of AVHs. Hoffman *et al.* (2008a) account for this by arguing that the middle temporal gyrus activation represents initial verbal content and/or prosody, which is then subsequently sent to the STG which creates the acoustic element of the AVH. Consistent with this, Shergill and colleagues (2004) found that it was only as individuals became aware of the AVH that activation spread to the bilateral STG and MTG. This led them to propose that 'activation in the temporal cortex [was] mainly occurring when the participant subsequently perceived auditory speech' (p. 517). Hoffman *et al.* (2008a) similarly concluded that 'more robust bilateral activation in the superior temporal gyrus arose somewhat later perhaps at hallucination onset – than the middle temporal gyrus activation' (p. 425). Overall, we are still in need of more large-scale studies (i.e. N > 20) of the activation preceding AVHs to form a reliable picture of the neural regions involved. However, we may tentatively suggest that involvement of inferior frontal and parahippocampal regions preceding AVHs suggests that inner

speech and memory-based processes, respectively, are involved in generating the raw material of AVHs, with activation of the STG then reflecting the perception of the AVH.

Summary of functional and structural neuroimaging studies

The most consistent findings that emerge from the above are: (1) an association between AVHs and structural abnormalities in the STG and IFG; (2) an association between AVH severity and hyperconnectivity in the AF; (3) functional activation during AVHs being most consistently seen in the IFG, STG, insula, cingulate, cerebellum and supramarginal gyrus; and (4) forms of activation/deactivation immediately preceding AVHs in the parahippocampal cortex, insula and IFG.

Lateralisation studies

The two largest fMRI studies of neural activation during AVHs discussed earlier in this chapter both found particularly the right IFG (including Broca's area) to be activated during AVHs (Diederen *et al.*, 2010; Sommer *et al.*, 2008). This suggests that laterality may be an important factor. Yet although Diederen, Sommer & Tendolkar (2011) in a pilot study found that language production was more lateralised to the right hemisphere in SZ:AVH+ compared to HCs, there was no difference in lateralisation between healthy voice-hearers and HCs. The authors thus suggest that lateralisation to the right in SZ:AVH+ is associated with psychosis per se, not AVHs. However, the frequency of AVHs in the SZ:AVH+ were significantly greater than the healthy voice-hearers. Hence, more work remains to be done to resolve this question.

If we turn away from changes in lateralisation in Broca's area to changes in lateralisation in temporal regions, the evidence for a link between unusual lateralisation and AVHs is stronger. Here we may first consider evidence from dichotic listening (DL) tasks. DL tasks are a non-invasive way to test hemispheric asymmetry and are also a test of the functional integrity of left temporal lobe language areas (Hugdahl, 1988). In this task, single auditory stimuli (such as consonant-vowel syllables, e.g. /ba/, /ta/, /ga/) are presented to both ears simultaneously, and the participant has to attempt to correctly identify the stimuli they heard, or if they heard two, which they heard best. It is typical for healthy individuals more often to report the stimuli reported to the right ear, as this information has direct access to the left temporal lobe, which is superior to the right temporal lobe in language-processing ability (ibid.).

In contrast, the stimuli presented to the left ear is initially sent to the right temporal lobe before having to be transferred across the corpus callosum to the left temporal lobe, which delays and attenuates the signal (ibid.). Using a DL procedure, several studies have shown that those with AVHs do not show the typical right ear advantage (e.g. Levitan, Ward & Catts, 1999). Recently, Løberg, Jørgensen & Hugdahl (2004) found that SZ:AVH+ showed the opposite pattern of ear advantage to HCs, showing a left ear, not a right ear advantage. Patients with a diagnosis of schizophrenia with a history of AVHs, but no current AVHs, did not show this left ear advantage. In addition to this finding, McKay, Headlam & Copolov (2000), as part of a comprehensive assessment of auditory processing between SZ:AVH+ and SZ:AVH−, found very few group differences, but did find that the former group were impaired compared to the latter on performance on a left monaural speech perception test, suggesting impaired right auditory cortex dysfunction or a problem in signalling between the auditory areas of the two hemispheres. This finding is supported by a recent study which (although using a global measure of hallucinations, rather than AVHs per se) found that in patients diagnosed with schizophrenia there was an association between hallucination severity and the extent to which the normal brain asymmetry in the temporal lobe (left greater than right), in terms of both volume and functional activation, was reduced (Oertel *et al.*, 2010).

Transcranial magnetic stimulation

Another methodology that allows us to investigate the neural activity associated with AVHs is transcranial magnetic stimulation (TMS). Whereas with electroconvulsive therapy,[2] best known to the public from the film *One Flew Over the Cuckoo's Nest*, a gross electric shock is simply applied to the patient's head, the induction of electrical current in the brain by TMS is precise, painless and done while one is wide awake. TMS works by holding a stimulator coil over the desired area of the scalp. A rapidly changing magnetic field in the coil induces an electrical current in the cortex directly below the coil, and can be aimed with an accuracy of around 1 cm. fMRI can be used to identity where a specific neural structure is in the individual participant, to ensure that the desired area is being stimulated. The electrical field induced in the cortex then causes the neurons in this area to fire. In slow TMS (done

[2] A systematic review by Tharyan & Adams (2005) suggested that ECT, when combined with antipsychotic medication, may prove beneficial for patients diagnosed with schizophrenia.

at a frequency of around 1Hz) this reduces the excitability of the region targeted. Given that some theories posit that AVHs are due to hyperactivation of speech perception areas, TMS can be used to try to reduce AVHs. At present, though, this technique is used as a last resort with medication-resistant patients (Aleman & Larøi, 2008).

Aleman, Sommer & Kahn (2007) in a meta-analysis of TMS for AVHs, found that TMS over the left TPJ region reduced the severity of AVHs (but not other positive symptoms) for some (but not all) patients. This suggests a key role for the left TPJ in the generation of AVHs in some patients. Since this meta-analysis, further studies have replicated the finding that TMS over the left TPJ region improves AVHs (Vercammen *et al.*, 2009). Some such studies have also used fMRI-guided TMS, in which fMRI is used to identify the location of brain activity during AVHs, then TMS is applied to this specific region. For example, Jardri *et al.* (2007) found that fMRI-guided TMS over the left TPJ cortex resulted in a 47 per cent improvement in an 11-year-old child's AVHs. This was a clinically significant improvement, and the child was able to return to school. We may ask what specific properties of AVHs, TMS actually improves. Horacek *et al.* (2007) found that the loudness, salience of the voices, as well as the amount of distress the voice-hearer experienced, were all imporved by TMS over the left TPJ. However, other facets, such as the frequency of voices, their reality, the number of voices, and the length of their utterances were not found to be significantly reduced by TMS.

Given that inferior frontal regions have been suggested to be involved in AVHs in the above review, what happens when TMS is conducted Broca's area? Although TMS over Broca's area has not been found to be beneficial for AVHs (Hoffman *et al.*, 2007; Schonfeldt-Lecuona *et al.*, 2004), it has been found that, for SZ: AVH+ with continuous AVHs, the greater the coupling between right Broca's area and Wernicke's area is, the less effective TMS over TPJ regions is, and, for SZ: AVH+ with intermittent AVHs, that the greater the amount of activation in Broca's area during AVHs, the less effective TMS over the TPJ is (Hoffman *et al.*, 2007). Again this starts to push us towards the need to posit the existence of sub-types of voices, which have different underpinning neural mechanisms, and may require different forms of treatment. This could be taken to suggest that whilst some AVHs are due to hyperactivity of the TPJ, others have their roots in connectivity problems between speech production and speech perception regions.

Aside from reducing cortical excitability, how does TMS have its effect? First, TMS may increase functional activity between areas of the brain that are 'disconnected' in patients with AVHs. For example,

TMS over the left TPJ results in an increase in connectivity between the left TPJ and the right insula in patients with AVHs (Vercammen *et al.*, 2010b). Second, at a cognitive level, which we will discuss further in the next chapter, improvements in source monitoring resulting have been found to result from TMS over the left TPJ (Brunelin *et al.*, 2006b). The meaning of this will be examined in the next chapter. It has also been found by Horacek *et al.* (2007) that TMS exerts its effects by both transcallosal as well as intrahemispheric connections, reducing brain metabolism in the left STG and its interconnected regions (e.g. hippocampus, insula), but increasing metabolism in the contralateral cortex and in the frontal lobes (e.g. the middle frontal gyrus). The potentially interhemispheric effects of TMS on language areas have also been noted by Andoh & Martinot (2008). They argue that TMS can cause functional reorganisation in homologous areas to where the 'virtual lesion' occurs, in order to compensate for the stimulated and disturbed area. Given the electrophysiological evidence for an involvement of the right TPJ (Line *et al.*, 1998; see electrophysiological section below) and right Broca's area (Sommer *et al.*, 2008) in AVHs, this may explain how left TPJ TMS can cause improvement in AVHs. Finally, it has also been proposed that TMS may exert its effect at a neurochemical level by inhibiting subcortical dopamine release (Aleman, Sommer & Kahn, 2007). It is such neurochemistry to which we now turn.

Neurotransmitters, antipsychotic drugs and AVHs

Insights into the proximal biological causes of AVHs can also be gained from studies of voice-hearers who have been treated with antipsychotic medication. Before we start building any model of AVHs based on the proposed mechanisms underpinning antipsychotic drug treatment, it is worth considering how effective these drugs are. Whilst estimates of their short- and long-term effectiveness vary, they do at least appear to be helpful for some voice-hearers (see Appendix A for further discussion).

Exactly what happens to people's AVHs when they take antipsychotic medication which has a useful effect? In an early study, Elkes & Elkes (1954) found that of the two patients in their sample who heard voices, chlorpromazine did not make the voices disappear, but only made the patients less bothered by them (i.e. the patients didn't shout and scream at their voices as much). For example, one patient stated that his voices 'did not worry him so much' (p. 563), and another who had been hearing a voice called 'Mr Knock', who put 'filthy thoughts into her mind', reported after chlorpromazine treatment that 'she did not bother

any more with Mr. Knock' as he 'did not annoy her so much' (ibid.). Fifty years later, Mizrahi *et al.* (2005), in a study of the effects of antipsychotic medication, found that before treatment patients thought that the medication would both eradicate the voices and help them be more detached from them. However, after six weeks the patients found that the drugs helped them be more detached from their symptoms – 'help deal, help stop thinking, and make the symptoms not bother' (p. 862) – but were less effective in taking away the voices altogether. Thus, while some patients find medication is able to eliminate their voices, for others the voices remain but are relegated to the backs of their mind, rather than vanishing (Kapur *et al.*, 2006).

A recent study by Schneider *et al.* (in press) examined how specific properties of voices changed over time in twenty-eight patients with a diagnosis of schizophrenia who were being treated with antipsychotic medication. This sample was a mix of first-episode admission and repeat admissions, and hostile and severely suspicious patients were excluded from the analyses. The outcome measure was the PSYRATS-AH, which assesses properties of voices including their frequency, duration, location, loudness, beliefs re-origin (i.e. caused by external events in the world versus being internal, self-related events), amount of negative content, degree of negative content, amount of distress, intensity of distress, disruption and voice-hearer control. Overall, 40 per cent of the patients stated that antipsychotic medication reduced their hallucinations. Four sub-scales of the PSYRATS-AH (using univariate tests) showed significant improvements, namely frequency (38% reduction in mean score), loudness (20% reduction), beliefs re-origin (42% reduction) and disruption (33% reduction). Reductions in loudness and disruption were found after 4 weeks, yet reductions in frequency and changes in the beliefs regarding the origin of the voice were found only after 10 weeks. Thus, there appears to be a two-stage effect of antipsychotics on voices: first a reduction in frequency and loudness, followed later by a reduction in the amount of disruption the voices caused and the patients' beliefs about the origin of the voices. Of course, given that there was no control group, we cannot necessarily attribute the causes of the change in the voices to drugs, as opposed to other factors associated with being treated.[3]

Biological mechanism of antipsychotic medication

Of the more than 100 antipsychotic medications available, all have the effect of decreasing dopamine levels (Kapur, 2003). In particular,

[3] See also Miller (1996) on changes in hallucinations following treatment.

D2 dopamine receptors are blockaded (Stahl, 2008). Hyperactivity of the mesolimbic dopamine pathway, which projects from the ventral tegmental area (VTA) in the brain stem to the nucleus accumbens in the ventral stratium (Stahl, 2008) has been proposed to be responsible for hallucinations (Weinberger, 1987), and particularly auditory hallucinations (Stahl, 2008).

Thus, if a pure D2 dopamine antagonist (i.e. a 'typical' antipsychotic such as chlorpromazine or haloperidol) is applied, this reduces the activation of the hyperactive D2 dopamine receptors in the mesolimbic dopamine pathway (Stahl, 2008). However, this action also reduces activity in other 'innocent' regions of the brain where there are D2 receptors. For example, D2 receptors working away in the nigrostriatal pathway also have their activity reduced and this causes as side-effects so-called extrapyramidal symptoms, which can include tremors and shaking and tardive dyskinesia (facial and tongue movements and grimacing) (ibid.).

Why should such over-activity of mesolimbic D2 receptors cause AVHs? At least two reasons exist. The first, as Stahl (ibid.) describes, is that there is a cortico-striatal-thalamic-cortical loop which creates a thalamic sensory filter controlling the amount of information that is sent to the cortex. In brief, the corticoacumbens glutamate pathway runs from the prefrontal cortex (PFC) to an area of the ventral striatum known as the nucleus accumbens. This in turn has inhibitory gamma-aminobutyric acid (GABA) connections to the thalamus. There is then a separate, direct route from the thalamus back to the prefrontal cortex, the thalamocortical pathway. This creates a circular loop between the thalamus and PFC which enables the amount of information that the thalamus sends to the PFC to be controlled in a form of 'sensory filter', which stops too much of the sensory traffic coming into the thalamus from escaping to the cortex, where it may overwhelm or confuse cortical information processing. In this loop, dopamine inhibits the GABA neurons which project to the thalamus. Thus, if dopamine levels increase here, the GABA neurons cannot inhibit the activity of the thalamus as effectively, leading to more information getting through the sensory filter of the thalamus. It has thus been proposed that if there is N-methyl-d-aspartate (NMDA) hypofunction in the VTA, this results in (1) mesolimbic dopamine hyperactivity, which reduces the ability of the GABA neurons running from the nucleus accumbens to the thalamus to inhibit the thalamus's activity, and (2) corticoacumbens glutamate pathway hypoactivity, which further reduces the firing of the GABA neurons in the nucleus accumbens, leading to further reduced inhibition of the thalamus. This, in theory, leads to the sensory filter

function of the thalamus breaking down, allowing more information to pass from the thalamus to the PFC, which may lead to AVHs (all taken from Stahl, 2008).

A second theory is the motivational salience hypothesis put forward by Kapur (2003). According to this hypothesis, excess dopamine makes internal representations of precepts, language and memories more salient. In particular, the mesolimbic dopamine system has been argued to be a critical component in the 'attribution of salience', a process in which 'events and thoughts come to grab attention, drive action, and influence goal-directed behavior because of their association with reward or punishment' (ibid., p. 14). Dopamine also adds flavour to the experience, making such experiences either desirable or aversive (Kapur, 2003). This theory is attractive, as it offers an explanation for the finding, described above, that patients' AVHs do not always vanish, but just bother them less.

But why should there be over-activity of D2 receptors in the first place?[4] As Stahl (2008) describes, going back a step to examine what normally influences the level of dopamine in the mesolimbic dopamine pathway, it has been found that the glutamate projection from the brain stem effects the mesolimbic dopamine pathway. The cortical brain stem glutamate connection normally acts as a brake on the mesolimbic dopamine pathway by an inhibitory GABA interneuron in the VTA. Hence, it is proposed that hypoactivity of NMDA receptors in this pathway means that the mesolimbic dopamine neurons are not inhibited and thus become hyperactive. This is all well and good but, as will be discussed in detail in Chapter 11, we need to go back further in the aetiological chain to the events happening in people's lives that cause dopamine levels potentially to change.

In contrast to the typical antipsychotics, atypical antipsychotics act both to block D2 receptors and serotonin 5HT2A receptors (Stahl, 2008). However, as atypical antipsychotics have been found to be no more effective than typical antipsychotics (Lewis & Lieberman, 2008), the additional serotonin action of these drugs may not tell us much about the mechanisms underlying AVHs.

Electrophysiological studies

Electrophysiological techniques are able to examine the electrical activity in the brain with a higher temporal resolution than fMRI, and can hence give us further insights into the neuroscience of AVHs. One

[4] In Chapter 11 we will go outside the head to look at events in the world that might cause such changes.

well-studied area is mismatch negativity. Normally when we hear a series of similar sounds and then a new sound appears, the brain reacts by generating an electrophysiological response called a mismatch negativity (MMN). This MMN involves automatic, pre-attentive detection of auditory changes (Näätänen, 1990). It has been proposed that the MMN arises by the novel stimulus being compared to our memory of the usual stimulus. If there is a discernible difference between the incoming auditory stimulus and the existing memory trace, the MMN is generated (ibid.). In a study of the MMN in SZ:AVH+ and SZ:AVH−, Fisher *et al.* (2008) report that reductions in frontal MMN amplitude are associated with clearer AVHs, and suggest that there is a relationship between AVHs and preconscious auditory stimulus detection. They also note two previous reports showing an association between pre-attentive processing and AHs, with MMNs in left frontal/temporal sites being negatively correlated with AH ratings (e.g. Hirayasu *et al.*, 1998). However, no MMN differences were found by van Lutterveld *et al.* (2010) in healthy voice-hearers. They did, though, find an increased P300 response in healthy voice-hearers. The P300 event-related potential is a positive electroencephalograph (EEG) deflection which occurs approximately 300ms after the presentation of an anomalous stimuli, and is thought to represent conscious processing of stimuli (Näätänen, 1990) and to be related to the degree of attentional resources that are deployed. Others have found reduced P300 activity over the left temporal lobe to be specific to SZ:AVH+ (Havermans *et al.*, 1999). Indeed, Papageorgiou *et al.* (2004) found that P300 amplitudes over the left temporoparietal region and at the left prefrontal area were lower in SZ:AVH+ before antipsychotic medication, as compared to after treatment.

What appears particularly promising are studies that have examined EEG coherence in SZ:AVH+ while they were hallucinating and while they were not. EEG coherence is basically a measure of how strongly two regions of the brain are talking to each other. High coherence indicates a high functional connectivity between these two areas, which may be either exciting or inhibiting each other (Sritharan *et al.*, 2005). They found that although coherence in the EEG alpha-band between Broca's and Wernicke's areas did not differ between voice-hearing and non-voice-hearing states, there was an increase in coherence between the right and left STG during AVHs. The authors suggest this may implicate the interhemispheric pathway between the auditory association areas in the two hemispheres in AVHs.[5]

[5] Much like Julian Jaynes (2000) proposed (Chapter 1).

Another very interesting finding is that in a study of eight SZ:AVH+, Line *et al.* (1998) found EEG activity one second before AVH onset occurring over the right temporoparietal region one second prior to the patient's report of AVH onset. The authors note that the right temporoparietal cortex is involved in the process of self-recognition, and since their study many have found this region is involved in such judgements (see Decety & Sommerville, 2003). For example, Keenan *et al.* (2001) found that during a Wada test (where one hemisphere is anaesthetised, the right hemisphere in this case) an individual was not able to recognise his own face, and thalamic-temporoparietal lesions have been found to impair the recognition of one's movements as one's own (Daprati *et al.*, 2000). Yet outside of AVH episodes, using probe tones Ford *et al.* (2009) found that SZ:AVH+ had a lower electrophysiological response in their *left* (but not right) primary auditory cortex than SZ:AVH−. They took this to indicate that a non-speech area of auditory cortex (Heschl's gyrus) was 'turned on' and 'tuned in' in voice-hearers in order to process internal acoustic information, and that this came at the cost of being able to process sounds in general (p. 65). The electrophysiological work of Ford and colleagues is discussed in more detail in relation to inner speech models in Chapter 9. For a good review of electrophysiological studies of hallucinations, the interested reader should consult van Lutterveld, Sommer & Ford (2011).

A particularly novel development arising from EEG studies in relation to AVHs comes from the study of microstates. These are transiently stable (*c.*100–200ms) distributed neural networks, which have been referred to as the 'atoms of thought', as they are thought to reflect specific conscious experiences (Lehmann & Koenig, 1997). Kindler *et al.* (2010) investigated such microstates in relation to specifically AVHs, looking at microstates in SZ:AVH+ when they were at rest. It was found that AVHs were associated with a shortening of a specific type of sub-second EEG microstate (class D, which has a fronto-central distribution). The authors suggest this microstate provides a protective cognitive function, and they speculatively propose that its shortened duration may impair the correction of errors involving misattributing self-generated inner speech to external sources. We will return to inner speech models in Chapter 9.

Genetics

What role might genetics play in AVHs? Given the areas found above to be involved in AVHs, we should expect that genetic research

into AVHs would target genes involved in these regions. Indeed, Hugdahl *et al.* (2008), in a review of the current state of knowledge of genetics and AVHs, suggest that genes involved in speech and language are likely to be implicated. The FOXP2 gene they note is a potential candidate, as patients with abnormal FOXP2 function show disturbed activation of language-related brain regions, such as underactivation of Broca's area (Liégeois *et al.*, 2003). They also observe that polymorphisms in FOXP2 have recently been associated with auditory hallucinations in patients diagnosed with schizophrenia (Sanjuan *et al.*, 2006). Another gene that Hugdahl and colleagues highlight is the cholecystokinin type A receptor (CCK-AR) gene, which impacts upon dopamine levels, and has been found to be related to persistent AVHs by Sanjuan *et al.* (2004). They also note Sun *et al.*'s (2005) finding of a gene (LMO4) being expressed differentially in left and right peri-Sylvian regions of the embryonic human brains, and speculate that given the laterality issues highlighted here in AVHs, this may play a role.

Thomas *et al.* (2007) found a correlation between the presence of AVHs in patients diagnosed with schizophrenia spectrum disorders and the presence of AVHs in a sibling of theirs who also had a schizophrenia spectrum disorder (77 pairs of twins in the USA). However, this was not found in a larger sample in India (136 pairs), which may have resulted from the overall higher prevalence of hearing voices in the American sample (83.4%) as compared to the Indian sample (64.3%). The authors take this to suggest a significant impact of non-shared environmental factors in the aetiology of hearing voices, but note that three other studies have failed to find correlations between AVHs in siblings who both have schizophrenia spectrum disorders (DeLisi *et al.*, 1987; Kendler *et al.*, 1997; Hwu *et al.*, 1997, as cited in Thomas *et al.*, 2007).

AVHs in other contexts have also started to be studied in terms of their genetics. In epilepsy research, Winawer *et al.* (2000) have linked the syndrome of autosomal dominant partial epilepsy with auditory features (ADPEAF) in families to an area of chromosome 10q. AVHs in Alzheimer's has been found to be associated with a specific genetic polymorphism of the 5-HT2A receptor polymorphism 102-T/C, with those with some form of the C102 allele being more likely to have AVHs (Holmes *et al.*, 1998).

Yet all voice-hearing genetic research is still in its infancy, and it appears most likely that epigenetic processes, interactions between genes, environment and psychological factors (see Chapter 11) will play a role in the genesis of AVHs.

Table 8.6. *Summary of neurophysiological findings*

Methodology	Finding
Neuroimaging studies	• Grey matter abnormalities in the STG and IFG • Hyperconnectivity in the arcuate fasciculus • Functional activation during AVHs in the IFG, STG, insula, cingulate, cerebellum and supramarginal gyrus • Activity immediately preceding AVHs in the parahippocampal cortex, insula and IFG
Lateralisation	• Signalling abnormalities between left and right temporal lobes, or right hemisphere dysfunction
Transcranial magnetic stimulation (TMS)	• Overactivity in temporoparietal junction • Role for connectivity between Broca's and Wernice's areas
Psychopharmacology	• Increased mesolimbic D2 dopamine activity
Electrophysiological	• Increase in coherence between right and left STG during AVHs • Increase in EEG activity one second before AVH onset over right temporoparietal region • Impaired communication (corollary discharge signal) between speech production and speech perception areas (see Chapter 9)

Building a neurobiological model of AVHs

How can we draw the findings of this chapter together? Table 8.6 summarises the key findings from this review. These may be integrated into a neuroanatomical model of AVHs, but before doing this it is worth noting the heterogeneity of the findings. How can we explain, for example, that some neuroimaging studies find Broca's area activation during AVHs, whilst others do not? We already noted sample size issues as a potential cause of this, but given that we identified two types of AVHs in Chapter 7, Type 1 (Dynamic), with its two sub-types, and Type 2 (Static), we must allow that these different types of AVHs (or indeed, other more fine-grained phenomenological distinctions) may have different neural underpinnings. This may explain some of the variability in the findings. Contrastingly, it may be that the same AVH can be caused by different neural pathways. These issues remain to be clarified.

One recent neuroanatomical model of AVHs is that of Allen *et al.* (2008). They propose that bottom-up dysfunction[6] occurs through

[6] I apologise for using the language of dysfunction as is common in this area. This is not meant to imply that voice-hearing is pathological. I am using the term in the sense of 'different to how non-voice-hearers' neural systems work' and it is not meant to imply judgement.

spontaneous hyperactivity in the STG which primes these areas, leading to 'over-perceptualization' (p. 187). This may result in increased bottom-up modulation from the auditory cortex to other cortical regions which let the person experience and perceive their own internal auditory activity in a more vivid sense. Such perceptions are likely to be felt as non-self-produced, they argue, due to a weakening of top-down control from ventral anterior cingulate, prefrontal, premotor and cerebellar cortices leading to poor self-monitoring and impaired experiences of agency. Areas involved in the regulation of emotion (e.g. parahippocampal and cingulated regions) then contribute to the affective valence of AVHs.

Allen *et al.* note that their model allows for 'a disconnection model, in which frontal regions fail to prime perceptual centres regarding the internal origin of self-generated speech' (p. 188) as well as for the possibility of reduced control by monitoring centres. Disconnection models of AVHs (e.g. Friston & Frith, 1995) propose that such experiences arise due to a failure of the connections between the frontal cortex and the temporal lobe.[7] Such models draw on studies like that of Muller-Preuss & Jurgens (1976, as cited in Friston & Frith, 1995), who found that specific cells in the auditory cortex of squirrel monkeys responded to externally produced sounds, but not to self-generated sounds from the monkey itself. Ploog (1979, as cited in Friston & Frith, 1995) concluded that the inhibition of these cells during self-produced vocalisation was caused by corollary discharge associated with vocalisation, possibly from the anterior cingulate cortex, which projects not only to Broca's area, but also to auditory areas. As applied to AVHs, this means such experiences could be understood as self-formed words and utterances that are experienced as externally produced. Yet Allen *et al.* (2008) also claim that 'we also hypothesise an increased activation or hypercoupling of speech production centres in the inferior frontal cortex and speech perception areas in left temporoparietal cortex' (p. 188). It is hard to see how one can simultaneously advocate a disconnection and hyperconnection model between frontal and temporal regions.

The neuroimaging findings reviewed here support Allen *et al.*'s proposal of involvement of the STG in AVHs. Both structural GM changes in the STG, its activation during AVHs, and its abnormal connectivity with language production areas suggest a key role for it in AVHs. However, Allen *et al.*'s model is unclear about what is the cause of the over-activity in the STG. At least two possibilities exist. First, AVHs

[7] We will see more evidence for such models in the next chapter.

could be conceived of as resulting from focal activity of an epileptic nature, originating solely in the STG. Indeed, as we saw at the start of this chapter, direct electrical cortical stimulation of the STG is sufficient for AVHs to occur (Penfield & Perot, 1963). Yet this seems unlikely to be the primary cause of AVHs, due to the variety of neural regions found to be associated with AVHs in this chapter. Indeed, the above findings support David's (1999) proposal that AVHs 'cannot be regarded as random "discharges" from a diseased brain, but rather as the distorted output of a complex cognitive system' (p. 95).

A more nuanced model, as suggested by Hunter *et al.* (2006) is that activity in the anterior cingulate cortex (ACC) (involved in attentional processes) drives spontaneous fluctuations of activity levels in the STG which result in AVHs. In support of this proposal, Hunter and colleagues found, using fMRI, that when non-voice-hearing individuals sat in silence, there were clear 'intermittent episodes of strikingly increased activity within speech-sensitive regions'. These regions included the left STG and the medial transverse temporal gyrus (the site of primary auditory cortex). These fluctuations were also accompanied by increases in activation in the anterior cingulate cortex, as well as in the right insula. Hunter and colleagues note that although such experiences did not result in AVHs in healthy individuals, a greater magnitude of such fluctuations could be the basis for AVHs.[8] The proposal for an involvement of the ACC in a similar way is supported by Fletcher *et al.* (1999), who observe that in addition to the STG in non-voice-hearing individuals being affected by activity in the PFC and the ACC individually, it is also sensitive to a combination of activity in the ACC and PFC above and beyond its sensitivity to activity of either region in isolation. They suggest that this reflects a modulatory effect of the ACC on prefrontal interactions with the superior temporal cortex. Indeed, the ACC has strong reciprocal connections with both the prefrontal and superior temporal cortices, which allow it to modulate the prefrontal inhibition of temporal regions (Fletcher *et al.*, 1999). In people at risk for development of schizophrenia, increased activation of the ACC and increased effective connectivity between this region and temporal and frontal areas has been found (Allen *et al.*, 2010). Allen and colleagues suggest that this is because it is attempting to compensate for a fronto-temporal system which was faulty, but had not yet failed (a failure which would result in the onset of AVHs). However, Hunter and colleagues' 'spontaneous fluctuation' model still needs to be able to account for the specific

[8] This will be discussed further in Chapter 10 as part of the default network.

phenomenology of AVHs as laid out in Chapter 4. For example, why should such spontaneous fluctuations so often result in commands or meaningful comments directed to the voice-hearer about ongoing events in their life?

In addition to the ACC, and in line with the evidence from this review, STG activation appears to result from temporally prior inputs from other neural areas such as the inferior frontal gyrus (via the AF) and/or the parahippocampal gyrus. Structural abnormalities in the IFG, in conjunction with its hyperconnectivity to the STG, may create abnormal activation in the STG, further kindled by structural abnormalities in the STG itself. Similarly, disinhibiton of the parahippocampal gyrus preceding AVHs may cause abnormal activation in the STG. The trend for studies to find STG/MTG activation increasing in the temporal lead-up to AVHs, with peak activation during the AVH itself, also suggests that activity in this area is caused by prior activation in other neural areas. The anterior cingulate may still play a role in modulating this network (see Chapter 10).

Whereas Allen *et al.* (2008) focus on a front-to-back disconnection, the evidence from EEG studies reviewed here suggests that side-to-side disconnection, specifically between the left and right STG, may play a role. This, when taken in conjunction with the lateralisation studies, and Vercammen *et al.*'s (2010a) work showing disconnectivity between regions such as the left TPJ and the right insula, suggests that a complex network of bilateral frontal and bilateral temporal regions may underpin AVHs. It is notable how both recent EEG and fMRI studies are finding more evidence of right hemisphere involvement.

Allen *et al.*'s (2008) model does not address (and was not intended to) the inter-relations between dopamine and neuroanatomical findings. Here we may consider Gray's (1998) theory. In this, structural abnormalities in the hippocampus, amygdala and temporal/frontal cortex cause hyperactivity in the mesolimbic dopamine pathway, which in turn causes a disruption to the integration of past experiences with current stimuli, which results in AVHs. In this model the limbic forebrain (e.g. prefrontal cortex, ACC), creates predictions of the upcoming state of the world from the person's current motor programmes, which it transmits to the nucleus accumbens via the subiculum (the inferior part of the hippocampus), which compares the expected state to the actual state. In voice-hearers it is proposed that the input from the hippocampus to the nucleus accumbens is disrupted leading, chemically, to hyperactivity in the mesolimbic dopamine pathway and, experientially, to novel, unexpected events, which are experienced as AVHs. Items which are unexpected are reactivated by feedback from the comparator system to those areas of the sensory neocortex (visual, auditory, somatosensory, etc.) in which they have just been non-consciously analysed. It is this reactivation by feedback from the comparator that selects these items for entry into consciousness. In this chapter we saw evidence of a role for the parahippocampal gyrus preceding AVHs. Gray (1995) argues that activity around this hippocampal region relates to its role as a 'novelty detector' that automatically draws attention when the organism is confronted with an unpredicted situation. Gray notes that the involvement of this area means that items that should be treated as 'expected/familiar' are in fact treated as 'unexpected/novel'. Others have focused specifically on the interaction between dopamine and cingulate regions, with Dolan *et al.* (1995) proposing that AVHs result from 'dysregulation in the dopaminergic modulation of cingulate neuronal activity with a resulting impairment in the functional integration of more remote, but anatomically connected, cortical regions' (p. 182).

The next chapters will attempt to translate these neuroanatomical findings into an explanation at the neurocognitive level. However, this will not be straightforward, given the multiple cognitive and affective functions of many grey matter structures in the brain. Such findings are open to multiple interpretations, and the meaning of the activation of these areas at the neurocognitive level is underdetermined. As Hein & Knight (2008) observe, 'the same brain region can support different cognitive operations depending on task-dependent network connections' (p. 2125). For example, although Allen *et al.* (2008) propose that the involvement of the parahippocampal gyrus relates to its role in emotional memory, it is also possible that its involvement could relate to attentional or self-monitoring processes. In order to assess what neurocognitive explanation is most suited to the current pattern of findings, it is necessary to assess whether leading neurocognitive models of AVHs are able to account for the findings of this review, and if not, then to consider how to potentially extend or revise these accounts based on these findings.

Chapter 8: summary of key points

- Spontaneous epileptic discharges in the STG are unlikely to be a good model for all AVHs (possibly accounting for Type 2, but not Type 1 AVHs).
- Neuroimaging studies typically show AVHs to be associated with grey matter abnormalities in the STG and the IFG, and hyperconnectivity in the AF.

- Functional activation during AVHs typically occurs in the IFG, STG, insula, cingulate, cerebellum and supramarginal gyrus, whilst activity immediately preceding AVHs is typically seen in the parahippocampal cortex, insula and IFG.
- TMS studies suggest that overactivity in the temporoparietal junction, as well as connectivity between Broca's and Wernice's areas, may be involved in AVHs.
- Increased dopamine activity may make certain cognitions more salient and play a key role in the generation of AVHs.
- Electophysiological studies show an increase in coherence between right and left auditory association areas during AVHs, as well as an increase in EEG activity one second before AVH onset over right temporoparietal region.
- A neuroanatomical model of AVHs can be built around impaired connectivity between frontal speech production areas and temporal/parietal regions involved in speech perception (with a potential modulatory role for the anterior cingulate), and impaired interhemispheric connectivity between auditory association areas.
- More research is needed into such neural underpinnings, however, using large samples and a variety of techniques, linking the specific phenomenological properties of AVHs to neural mechanisms.
- There is a need to build neurocognitive models to help us better understand AVHs, a task which we turn to next.

9 Neuropsychological models I: inner speech

Attempting to deduce what causes AVHs simply from neurological findings is problematic, not least because of the multiple functions of each region of the brain. There is hence the need for an account at the neuropsychological level because, as Churchland (1986) has argued, 'neuroscience needs psychology because it needs to know what the system does' (p. 373). Similarly, Coltheart & Langdon (1998) argue that 'it can be very hard to understand what a system is actually doing if one's only information about it is a description at the physical-instantiation level. A description at the abstract-theory level will be far more enlightening' (p. 150). Thus, co-activation of, for example, Wernicke's and Broca's areas of the brain in AVHs is relatively uninformative unless we know what cognitive functions these areas are involved in. Furthermore, as most areas of the brain are involved in many possible tasks (i.e. the meaning of their activation is underdetermined), it makes sense to be guided by an explanation of AVHs which proposes what normal cognitive processes have gone awry to produce them.[1] Yet this can be a two-way process, as functional neuroimaging can be used to inform and challenge a model of AVHs which was initially conceived in pure cognitive psychological terms (Buchsbaum & D'Esposito, 2008). For example, if cognitive models of AVHs had no mention of memory processes, and neuroimaging studies showed neural areas associated with memory to light up during AVHs, then cognitive models would need to be updated to take into account the likely involvement of memory processes.

Frogs and salt

If one rules out the theory that AVH comes from an external, ontologically independent, supernatural being, then we are left with AVHs having

[1] Or, less negatively, potentially to understand what functional purpose AVHs might be providing.