# Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex

Heledd C. Hart
Alan R. Palmer
Deborah A. Hall

MRC Institute of Hearing Research
University Park
Nottingham
NG7 2RD
UK

Tel : (+44) 115 922 3431
Fax : (+44) 115 951 8503
e-mail : heledd@ihr.mrc.ac.uk

**Abstract**

Hall et al. (2002) recently showed that pulsed frequency-modulated tones generate considerably higher activation than their unmodulated counterparts in non-primary auditory regions immediately posterior and lateral to Heschl's gyrus (HG). Here, we use fMRI to explore the type of modulation necessary to evoke such differential activation. Carrier signals were a single tone and a harmonic-complex tone, with a 300-Hz fundamental, that were modulated at a rate of 5 Hz either in frequency, or in amplitude, to create 6 stimulus conditions (unmodulated, FM, AM). Relative to the silent baseline, the modulated tones, in particular, activated widespread regions of the auditory cortex bilaterally along the supra-temporal plane. When compared with the static tones, both AM and FM tones generated significantly greater activation in lateral HG and the planum temporale; replicating Hall et al. (2002). These activation patterns were largely overlapping, but direct comparisons between AM and FM revealed greater activation by the variation in amplitude than in frequency, plus a small part of the posterolateral region in the right hemisphere whose response was specifically AM dependent. The dominant pattern of activation was that of co-localised activation by AM and FM which is consistent with a common neural code for AM and FM within these brain regions.

**Introduction**

Communication signals and other everyday sounds change in frequency and amplitude on a moment-to-moment basis. As a first step in understanding the way in which the auditory system processes such spectrally and temporally complex sounds, responses have been studied to simpler stimuli which allow selective manipulation of specific features of the acoustical waveform. Convenient stimuli for studying the effect of changing the temporal characteristics of sounds are frequency and amplitude modulated single and complex tones. Sinusoidally modulating either the frequency or amplitude of a sound generates a well controlled stimulus in which the temporal characteristics are determined by the modulating waveform.

Previous functional magnetic resonance imaging (fMRI) studies have investigated brain responses to acoustical signals that are sinusoidally modulated in frequency (Hall et. al., 2002) or amplitude (Giraud et al., 2000) by contrasting them against their unmodulated counterparts. Hall et al. (2002) used single-frequency and harmonic-complex tones. Activation by these frequency-modulated tones was greater than by the unmodulated tones in bilateral Heschl's gyrus (HG), anterolateral and posterolateral parts of superior temporal gyrus (STG), and superior temporal sulcus (STS). An especially high response to the frequency-modulated harmonic-complex tone was reported in the posterolateral region. Giraud et al. (2000) also found stronger auditory activation to amplitude modulated white noise than to unmodulated noise. This differential activation was located bilaterally in primary and non-primary areas, but was strongest in posterolateral regions of the STG and STS.

Neuroimaging studies have investigated the effects of temporal variation by using other types of sounds that are not sinusoidally modulated. Again, their findings converge on the importance of non-primary auditory activation in the analysis of

3

temporally-varying sounds. Binder et al. (2000) used fMRI to study responses to the effects of spectral variation using sequences of single-frequency tones that stepped in frequency by at least 10 Hz, every 666 ms (i.e. at a rate of 1.5 Hz). Non-primary posterolateral areas were more strongly activated by the tones whose frequency changed over time than to unmodulated white noise.

The following studies using positron emission tomography (PET) also indicate an important role for bilateral posterolateral non-primary auditory regions in the processing of temporally varying sounds. Thivard et al. (2000) studied cortical responses to synthetic sounds that were similar to vocal sounds in structure, with spectral maxima that were modulated in time. Relative to matched signals that had a stationary spectral profile, spectrally varying sounds generated bilateral auditory cortical activation that was posterior to the lateral extension of HG. Griffiths et al. (1998) manipulated the time structure in sounds by presenting a sequence of binaural pitches that either had tuneful pitch variations (melodies) or no melody. An effect of melody was reported in bilateral non-primary auditory areas in the posterolateral and anterior parts of the STG. Griffiths et al. hypothesize that these areas are involved in the general analysis of sound sequences. Zatorre and Belin (2001) have studied the response to spectral and temporal variation by systematically manipulating the frequency and timing information in a sequence of single-frequency tones. Spectral variation involved tone sequences that varied in the spacing of their frequency components (from 1 octave to 1/32 octave frequency separation), but with a fixed rate of change between tones. Temporal variation involved tone sequences that varied in the rate of change (from 1.5 Hz to 48 Hz) between two fixed frequencies. This parametric study is conceptually different from the studies described so far (Giraud et al., 2000; Hall et al., 2002; Thivard et al., 2000) in that there is no unmodulated

baseline stimulus and so spectral and temporal conditions must be compared against one another. When temporal conditions were subtracted from spectral conditions, a greater response to spectral changes was found in bilateral non-primary auditory areas, but in anterior, not posterior, STG.

Physiological and psychophysical methods have been used to investigate whether the basis of AM and FM processing involves neural mechanisms that are separate or shared. Both alternatives have received empirical support. "Selective adaptation" studies have provided evidence that AM and FM stimuli are processed independently by parallel pathways in the auditory system (Kay, 1982; Regan and Tansley, 1979). These studies show that exposure to FM tones elevates the psychophysical detection threshold of FM, but not AM, tones. Similarly, exposure to AM increases the AM detection threshold, without affecting the FM threshold. Based on this evidence, and the fact that FM can easily be distinguished from AM at most modulation rates, the authors postulate separate processing channels for AM and FM. However, Wakefield and Viemeister (1984) have questioned the interpretation of these adaptation studies and suggest that the observed effects reflect changes in the decision strategy of the subjects rather than the existence of AM and FM specific channels.

Mäkelä et al. (1987) measured auditory cortical evoked magnetic fields to AM and FM tones, presented in pairs separated by 500 ms in four different combinations (FM-AM, FM-FM, AM-FM, AM-AM). They found that the decrease in response amplitude from the first to the second response was significantly smaller for pairs with non-identical stimuli, and thus suggest that the neuronal populations activated by AM and FM differ. They report only very small differences in the observed source locations for FM and AM, indicating that the activated cell populations partly overlap

or intermingle; or that the differentiation of responses is already made at more peripheral stations of the auditory pathway.

Different responses to FM and AM tones have been observed within single cells of the cochlear nucleus and inferior colliculus of the rat (Møller, 1971; Rees and Møller, 1983). Physiological data about the specificity of auditory cortical responses to FM and AM is less well understood. In the auditory cortex of awake cat, some neurones respond more strongly to changing stimulus frequencies than to constant frequency sounds (Whitfield and Evans, 1965). In anaesthetised cats, a spatial segregation of units responsive to the rate of FM has been measured along the isofrequency axes of the primary auditory cortex, but the distribution was variable across cats (Heil et al., 1992; Mendelson et al., 1993). In monkey auditory cortex, some neurones respond selectively to AM sounds presented at 2-32 Hz (Urbas et al., 1986). Analogous neuronal populations in human auditory cortex could form a basis for psychophysical channels selectively sensitive to changes of frequency and amplitude.

In contrast, other psychophysical and physiological studies have supported the theory that AM and FM are processed by a common neural code. Saberi and Hafter (1995) postulate that FM information is transduced at a relatively early stage in the auditory system so that a frequency change is transmitted as a change in amplitude, resulting in both AM and FM coding occurring by phase-locking of auditory afferents to the modulation envelopes. This view is based on the ability of human subjects to perceive a single image in the head depending on the phase difference between an FM sound presented to one ear and an AM sound with the same modulation rate and carrier to the other ear. Observers were able to accurately localise the position of this intracranial image, suggesting that the cues provided by these sounds were not

affected by use of a different type of modulation at the two ears and therefore that coding for the two modulation types is similar, at least at neural relays above the brainstem where binaural convergence takes place. Other evidence suggesting a common mechanism for modulation rate comes from neurophysiological data obtained from rats, cats and marmosets showing that best modulation frequencies of single primary auditory cortical units are significantly correlated for AM and FM (Eggermont, 1994; Gaese and Otswald, 1995; Liang et al., 2002). Liang et al. (2002) report that, in the primary auditory cortex of awake marmosets, individual cortical neurons, as well as populations of neurons, have common features in their discharge rate and synchrony-based responses to both sinusoidal AM and FM.

Imaging studies indicate that sounds which vary in frequency and in amplitude over time *both* activate non-primary auditory cortex when compared to unmodulated stimuli, particularly in the posterolateral region (Giraud et al., 2000; Griffiths et al., 1998; Hall et al., 2002; Thivard et al., 2000). However, the hypothesis that this brain region is involved in processing both types of modulation has never been directly tested in an imaging study. Some studies have actually confounded AM within the FM stimulus by using a sequence of tone bursts that were pulsed on and off during the stimulation (Binder et al., 2000; Hall et al., 2002). Studies that specifically seek to segregate responses to AM and FM must therefore use non-pulsed, continuous stimuli to avoid stimuli that have both AM and FM components.

The auditory cortex responds to many different stimulus features which must be carefully controlled in subtraction paradigms. It is important that responses attributed to the AM or FM components of the stimulus are actually due to the modulation and not due to some other stimulus characteristic. The rate of modulation is an important factor in determining the response, especially its shape. For low rates

of amplitude variation (e.g 1.5 Hz), the fMRI response shows a sustained response over the stimulation period, but higher rates of variation (e.g. 48 Hz) evoke a transient response to the stimulus onset that decays over the period of stimulation (Giraud et al., 2000; Harms and Melcher, 2002). Sound level is an additional factor that determines the magnitude and extent of the auditory response (Brechman et al., 2002; Hall et al., 2001). fMRI evidence indicates that for conservative comparisons between different types of sounds, the component levels of those sounds should be matched for loudness, rather than root mean square power (Hall et al., 2001). The carrier signal can also influence the robustness of the activation. For example, Hall et al. (2002) showed that the effect of FM was more reliably seen using a harmonic-complex carrier, compared to a single frequency carrier, but was nevertheless significantly present for both carrier signals. Thus, a simple way to control for the characteristics of the carrier is to ensure that the effects of AM and FM generalise across different carrier signals.

In the present study, we use 3 T fMRI to investigate whether or not the response to modulation is dependent on the characteristic of modulation (i.e., frequency or amplitude), for carrier signals presented at the same rate and sound level. An issue of particular interest was whether there is any spatial segregation of cortical areas responsive to FM and AM, or whether the same auditory areas are activated by temporally-modulated sounds compared to unmodulated sounds irrespective of the type of modulation.

**Materials and Methods**

**Stimuli**

The stimuli were synthesised at a sampling frequency of 48 kHz using two carrier tones; a 300-Hz single tone and a harmonic-complex tone. The harmonic-complex tone had a 300-Hz fundamental and consisted of four harmonics at 300, 600, 900 and 1200 Hz. 300 Hz was chosen to be distant from the frequency of peak energy of our scanner noise (1.92 kHz, plus higher harmonics; Hall et al., 2000). Crossing carrier signal (single frequency and harmonic complex) by modulation (unmodulated, AM, FM) generated six stimulus conditions, plus a silent baseline condition (see Figure 1(i)). Each tone was presented continuously for 6.8s, rather than pulsed, so that the AM was not introduced into the FM stimuli. The carrier signal, x(t), was either not modulated;

$$x(t) = A_c \sin(2\pi f_c t)$$

frequency modulated sinusoidally at a rate $f_m$ of 5 Hz and a depth $\beta$ of 10 (50 Hz)

$$x(t) = A_c \sin(2\pi f_c t + \beta \cos(2\pi f_m t))$$

or amplitude modulated sinusoidally at a rate $f_m$ of 5 Hz and a depth $m$ of 1 (100%)

$$x(t) = A_c \sin(2\pi f_c t) (1 + m \cos(2\pi f_m t))$$

where $A_c$ and $f_c$ are the amplitude and frequency, respectively, of the carrier component. Modulation depths were chosen to produce highly perceptible AM and FM that had been shown, in previous studies, to generate robust auditory activation (Giraud et al., 2000; Hall et al., 2002). For AM, m = 1 was used (100% AM). An FM modulation depth of 50 Hz (where ß = 10; the modulation depth divided by the modulation rate) was the same as that used by Hall et al. (2002). For harmonic-complex tones, use of ß = 10 prevents the upper sidebands of one harmonic from overlapping with the lower sidebands of the next harmonic. The AM and FM stimuli

were informally matched by an experienced listener to be approximately matched for modulation depth. The principal frequency components and the sidebands generated by modulation are shown in Figure 1(i) for each stimulus. For AM, the frequency spectrum contains three components: the principle component is equal to $f_c$ and two adjacent half-amplitude sidebands that occur at a distance equal to that of the $f_m$. Thus, for the 300-Hz tone, amplitude modulated at a rate of 5 Hz, the components are 295 , 300 and 305 Hz. For FM, the spectrum has more sidebands (each separated by $f_m$) which again decrease in amplitude away from $f_c$. For 300-Hz FM tone with a modulation depth of 50 Hz, the principal component is 300 Hz plus 19 components on either side of the carrier. A 5-Hz modulation rate was specifically chosen because the auditory cortex generates a sustained response to stimuli presented at a rate of 4-8 Hz (Harms and Melcher, 2002; Giraud et al., 2000). In addition, low modulation frequencies have an important psychological validity, being the most crucial for speech recognition (e.g., Houtgast and Steenekenen, 1985).

*** Figure 1 ***

The stimuli were generated at a loudness equal to that of the 300-Hz tone at 96 dB SPL and were matched using a computational model of loudness summation (Moore et al., 1997; see also Moore et al., 1999). The overall power of each stimulus at the ear was measured by mounting the fMRI headphone system on KEMAR (Burkhard and Sachs, 1975), equipped with a Brüel and Kjær microphone and measuring amplifier. These measurements revealed that the presentation level of the single tones was 96 dB SPL and the harmonic-complex tones was 83 dB SPL (+/- 0.5dB).

**fMRI acquisition**

Functional imaging was performed in twelve normally-hearing subjects (7 female and 5 males), whose mean age was 26 years. The study was approved by the Nottingham University Medical School Ethics Committee and was undertaken with the understanding and written consent of each subject. We used a gradient echo EPI sequence (TR = 9s, TE = 24 ms) on a 3 Tesla scanner with a head gradient coil set and a transmit/receive volume head coil (Nova Medical Inc, Massachusetts). The voxel size was 3 mm isotropic. Sets of brain images were acquired at the transitions between stimulus conditions using a clustered volume acquisition protocol, such that all 32 brain slices were acquired in 2.2 s (Figure 1(ii)). This sparse imaging method reduces the interference on auditory cortical activation from the intense background scanner noise (Hall et al., 1999). Furthermore, a pilot study using these stimuli, demonstrated that the 6.8 s silent period between volume acquisitions (TR = 9 s) materially improved the sensitivity of detecting auditory cortical activation compared with an acquisition protocol in which the silent interval was 4.2 s (TR = 6 s). The first two subjects to be scanned performed a gap detection task to maintain their arousal and attention to the stimuli. Each 6.8 s stimulus block contained a gap at a random position and subjects were asked to press a button as soon as they heard these gaps. After concern that this paradigm introduces additional amplitude changes within the stimuli, the remaining ten subjects were scanned using the following alternative paradigm. In each condition, tones were presented continuously for 6.8 s and conditions were presented in a random order. To maintain arousal, subjects were required to identify whether the sound was static (constant) or modulated (fluctuating) by pressing one of two buttons at the end of each stimulus condition. Each stimulus

11

condition was repeated 40 times, distributed across two sequential imaging runs. The total experimental time was 42 minutes.

A T1-weighted EPI image (3x3x3 mm resolution) was also acquired for each subject in the same imaging session. This type of image shows the same tissue distortions as the functional image, but provides enhanced contrast between grey/white matter and CSF, thus enabling visualisation of individual auditory cortical morphology.

**Data analysis**

For each subject, image data were motion corrected across the two imaging runs, transformed into a standard brain space using template matching, and spatially smoothed with a Gaussian filter of 6 mm, using SPM99 software (www.fil.ion.ucl.ac.uk/spm). Analysis of activation was also performed with SPM99, using a general linear model (Worsley and Friston, 1995). There was no observable difference between the activation patterns for the two different paradigms and data for all 12 subjects were combined in a random-effects approach using a two-stage analysis procedure (Holmes and Friston, 1998). Random-effects analyses are essential for valid population inference about the *typical* effects in fMRI datasets where the inter-scan variability within an imaging session is small in comparison to the variability of responses from subject to subject, as is generally the case in fMRI (McGonigle et al., 2000). Stimulus effects were conducted separately for all subjects, using a first-level, fixed-effects analysis that computed the within-session error variance (residual scan-to-scan variability) as the only component of variance. This first stage used a design matrix that comprised 7 variables, one for each stimulus condition. Each variable defined which set of images constituted that particular

stimulus condition using a simple vector of 1s (stimulus on) and 0s (stimulus off). Statistical contrasts between stimulus conditions were specified by a linear combination of these variables. The outputs of these individual analyses were then entered into a second level which tested the significance of the statistical contrast by assessing the between-session component of variance.

Firstly, we conducted a preliminary examination of the data to determine the typical pattern of auditory-evoked activation by contrasting all 6 tones against the silent baseline. For this non-hypothesis-driven analysis, we set a conservative probability threshold of $P<0.05$, correcting for multiple comparisons. Subsequent analyses were performed to determine the following specific issues : 1) tonotopic differences evoked by the spectral difference between the two carrier tones, matched for pitch, 2) whether there was any topographic segregation of auditory cortical areas responsive to AM and FM, presented at the same modulation rate and, 3) whether response magnitude was determined by the type of modulation and in particular, which type of modulation was necessary to evoke the higher activation of non-primary auditory areas observed by Hall et al. (2002). For these analyses, we set the probability threshold to $P < 0.001$ and did not correct for the number of separate multiple comparisons across the entire brain since we were testing a set of specific hypotheses about auditory cortical function.

**Results**

Subjects were able to discriminate unmodulated from modulated tones with a high degree of accuracy (mean score = 96%, stdev 5%). Data for the two subjects scanned whilst performing the gap detection task were comparable (mean score 96%, stdev 5%).

The contrast of all tones against silence revealed widespread bilateral auditory activation along the lateral two thirds of HG and surrounding areas of the STG. Additionally, small clusters of bilateral activation were focussed around the central sulcus (36 and 16 voxels in the left and right hemispheres respectively). The peaks of these clusters occurred at $x = -39$, $y = -24$, $z = 69$ mm in the left hemisphere, and $x = 36$, $y = -36$, $z = 72$ mm in the right. These peaks coincide with the location of sensorimotor cortex. We therefore suggest that these activations were task related and were likely to be evoked by the subjects' button-press response at the end of each stimulus condition.

Figure 2 shows the distribution of activation along the STP for separate tone conditions relative to the silent baseline (P < 0.001, uncorrected for multiple comparisons). Activation patterns varied across subjects, but the three subjects presented here are representative of the range observed. In general, the unmodulated 300-Hz tone and the unmodulated harmonic-complex tone yielded little or no significant sound-evoked activation. The unmodulated single-frequency tone produced auditory cortical activation in four out of the twelve subjects. The unmodulated harmonic-complex tone produced activation in five out of twelve subjects. In general, when activation was present, it was located on HG and/or in auditory cortical areas posterior and lateral to HG. Both types of modulation generated bilateral auditory activation that incorporated parts of HG and surrounding non-primary areas. Auditory cortical activation was observed for eleven out of twelve subjects for the FM single tone, and for all twelve subjects for the FM harmonic-complex tone, the AM single tone and the AM harmonic-complex tone. For these

sounds, activations were observed in medial HG, which includes the putative site of the primary auditory area, and in surrounding non-primary areas.

*** Figure 2 ***

### 1) Differences between harmonic-complex and 300-Hz tones

General effects of the carrier signal were assessed directly by computing the difference between the harmonic-complex and the 300-Hz tones, for each subject, combined across the static, AM and FM conditions ($P < 0.001$). Five subjects showed no significant differential activation. For the remaining subjects, the greater activation by the harmonic-complex tones was extremely variable in its extent and location. For example, the size of the activation clusters ranged from 1 to 62 voxels; in some subjects occurring only in the left hemisphere, and in others occurring only in the right. When activation was present, it generally occurred in HG and auditory cortical areas immediately posterior to HG. The contrasts for the individual subjects were then entered into a random-effects group analysis. This group analysis revealed no significant difference between carrier signals ($P < 0.001$). This lack of significant difference between the activations observed for the two carrier signals is probably due to the high degree of inter-subject variability observed and the conservative nature of the analysis.

### 2) Topographical segregation of AM and FM responses

If the processing of AM and FM occurs in different auditory regions then the activation patterns should be spatially segregated. The degree of segregation was tested in a number of ways. First, we generated a thresholded statistical map of activation for each type of modulation relative to the unmodulated stimuli (P<0.001

uncorrected for multiple comparisons) and report the topographical organization of the two resulting activation maps. Second, we directly computed the differential activation by AM relative to FM and FM relative to AM (again P<0.001 uncorrected for multiple comparisons). Finally, we used unthresholded maps of the standardised magnitudes (the beta parameters for each stimulus condition extracted from the general linear model) to generate 2-D plots of the spatial distribution of response magnitudes.

*** Figure 3 ***

First, the effect of modulation was tested by contrasting each modulated tone against its unmodulated counterpart. The contrast between modulated and unmodulated stimuli was computed for each subject, for the harmonic-complex and single-frequency tones combined. These contrasts were then entered into a random-effects group analysis. Analyses were computed separately for the AM and FM conditions. The results shown in Figure 3 demonstrate that both types of modulation generated significantly greater auditory cortical activation than did the unmodulated tones. The activation maps overlapped one another by approximately 55%. Common areas of activation included HG and a region immediately posterior to it, with this posterior region extending towards the lateral convexity of the STG. Activation by AM involved almost 20% more voxels than the activation by FM and it generally extended to the lateral convexity in the region posterior to HG, as well as ventrally towards the STS. The activation by FM included a small portion of left medial HG (the site of the putative primary auditory cortex), plus a region on the posterior border of right HG which undercut HG itself.

We then directly computed the differential activation by AM relative to FM ('AM-FM'). Results are shown by the areas outlined in black in Figure 3. The 'AM-FM' contrast identifies brain areas that are activated i) both by AM and FM, but with AM>FM and ii) exclusively by AM. These effects were shown by a small number of voxels, which were located in non-primary auditory cortex bilaterally. Brain areas where the 'AM-FM' activation map overlapped with both the 'FM-unmodulated' and the 'AM-unmodulated' maps indicated those areas of co-activation, where AM>FM. These areas can be seen in Figure 3 as those which are both outlined in black and coloured yellow. On the right planum temporale, 3 voxels with a peak at 60 –12 3 mm and 1 voxel at 54 –27 6 mm showed this pattern of co-activation. Brain areas where the 'AM-FM' activation map overlapped only with the 'AM-unmodulated' map indicated those areas where AM had a specific effect. A single area met the criterion for being specifically AM-dependent and can be found in Figure 3, outlined in black and coloured green. This area comprises 3 voxels in the right planum temporale posterior and lateral to the main axis of Heschl's gyrus, with a peak at 63 –12 3 mm. Relative to the size of the overlapping activation between 'AM-unmodulated' and 'FM-unmodulated' (83 voxels in the left and 88 voxels in the right hemisphere, shown in yellow in Figure 3), the AM-specific area is very small. The 'AM-FM' activation map identified a further 3 voxels in the left planum temporale (with a peak at –63 –42 9 mm), but this region was not activated by 'AM-unmodulated', and so the AM-dependent response is different from that shown by the right planum temporale. We also directly computed the differential activation by FM relative to AM, but the effects did not reach significance, suggesting that no brain areas are differentially activated by FM.

*** Figure 4 ***

Second, the spatial distribution of the activation by modulation was mapped using the model response estimates for these conditions. Bilateral windows were specified that incorporated voxels from $|39|$ to $|66|$ mm in the X dimension and -45 to +12 mm in the Y dimension to include the entire axis of HG plus the activated posterior and lateral auditory regions. The plane of section included regions along the supratemporal plane, and thus was angled by 34º using the T1-weighted brain image for the group. The beta parameters from the random-effects analysis reflect a standardised magnitude of each stimulus effect, and these were extracted for all voxels within the windows. Beta parameters are derived from the general linear model implemented in SPM99 and correspond to the model predictions of the magnitude of the condition-specific activation. We assume that the standard errors differ little across the region of interest because the data have been spatially smoothed. The plots revealed a spatially coherent ridge (peak value = 1.64 for AM harmonic-complex, at $x = 64$, $y = -12$, $z = -12$ mm) in a region posterolateral to HG that was present in the two hemispheres for both AM and FM tones (Figure 4(i)). The ridge represents an elevated response, whose peak was focused on the lateral part of the STG, but also incorporated lateral HG. Its spatial location was highly similar for both types of modulation, although its magnitude appeared greater for AM than FM stimuli. In contrast, the magnitudes for the static harmonic-complex tone did not exceed the range -0.34 to 0.48 and the plots revealed a 'flat landscape' across the STP.

In summary, although AM activated a somewhat larger region of the auditory cortex than FM and extended to the lateral edge of the STG, there was no clear segregation of the foci (peaks) of activated cortex between the two types of modulated

18

stimuli. As a final point, we draw attention to the correspondence between the posterolateral location of the high response to modulation obtained here and that obtained in a previous study (Hall et al., 2002) (Figure 4(ii)). Thus, the present results replicate the previous finding that an especially high response to a spectro-temporally modulated tone occurs in a posterolateral region of the non-primary auditory cortex.

### 3) *Effect of modulation on response magnitude*

The activation maps shown in Figure 3 and the plots of the response magnitude estimates shown in Figure 4(i) suggest that the AM tone elicits greater activation than the FM tone. This hypothesis was statistically tested in two ways. First, we measured the condition-specific effects on the mean signal magnitude within a region of activation that showed a significant response (P<0.001) to *both* amplitude-*and* frequency-modulated tones relative to the unmodulated tones. In a second, more stringent, test we contrasted the AM against the FM conditions in a random-effects group analysis for each and every voxel in the STG.

The defined region of activation used in the first analysis corresponds to the yellow region presented in Figure 3 and incorporates part of the posterolateral non-primary auditory region. For this region, a mean time series was extracted for each subject using the unsmoothed image data. Time series were corrected for any low-frequency signal drift up to a maximum frequency of 0.46 cycles per minute and, for each subject, the mean % signal change relative to the silent baseline was used as a normalised measure of each condition-specific effect (Figure 5). An analysis of variance revealed that the AM tones generated significantly greater response magnitude than the FM tones ($F$[1, 115]=13.4, P<0.001).

\*\*\* Figure 5 \*\*\*

The random-effects analysis revealed a greater response to AM in bilateral auditory cortex (P < 0.001). These effects were small at this threshold, being 3 voxels in size on the left and 4 voxels on the right, but on both sides were located in lateral non-primary cortex. The differential activation was located at $x = -63$, $y = -42$, $z = 9$ mm and $x = 60$, $y = -12$, $z = 3$ mm respectively (P<0.001). In the right hemisphere, this activation was located within the yellow shaded region in Figure 3. These results indicate that, while both types of modulation strongly activated lateral parts of HG and posterolateral regions, greater activation was evoked by the variation in amplitude, than in frequency. Moreover this differential activation was focused around those non-primary lateral auditory areas close to the lateral convexity of the STP on or behind HG.

**Discussion**

The present study revealed widespread activation in primary and non-primary auditory cortical fields in response to AM and FM stimuli. While both types of modulation strongly activated lateral parts of HG and posterolateral regions, greater activation was evoked by the variation in amplitude, than in frequency. The overriding pattern of activation was that of co-localised activation by AM and FM, although a small part of the posterolateral region in the right hemisphere responded specifically to AM, and not to FM.

### *Lack of activation by unmodulated tones*

The present study revealed little or no activation to unmodulated stimuli relative to the silent baseline for more than half of the subjects scanned (see Results). Previous fMRI studies have revealed robust auditory cortical activation for unmodulated 300-Hz tones, using a comparable stimulus duration and sound presentation level (Hall et al., 2002; Hart et al., 2002). A key methodological difference between these two studies and the present study is the fact that, here, the stimuli were presented continuously, rather than pulsed on and off. Presenting the stimuli as a train of sounds introduces temporal variation into the stimulus epoch. It has been shown that temporally-varying stimuli produce more auditory cortical activation than do those that are either unmodulated (Hall et al., 2002) or have a slower temporal variation (Zatorre and Belin, 2001). The presence of temporal variation within the sound may therefore facilitate the detection of activation by unmodulated tones. Without any temporal variation, unmodulated tones may evoke only a small change in the fMRI response, which in some subjects does not exceed the statistical threshold for detection.

### *Effect of the harmonic structure in the sound*

When the harmonic-complex and the 300-Hz tones were directly compared in a random-effects group analysis (for unmodulated and modulated conditions combined), no significant difference was found between carrier signals. This finding is in contrast to data presented by Hall et al. (2002), who showed that, as spectral complexity was increased, by adding five harmonics to a single frequency tone, activation increased in primary and non-primary auditory cortex. Hall and her colleagues found a high degree of overlap of the areas responsive to spectral and

21

temporal complexity, but activation induced by the temporal cues was stronger than that induced by spectral cues, as indicated by higher magnitude signal differences and T scores. In the present study, for each condition (unmodulated, FM, AM), the mean magnitude of the signal (Figure 5) was always greater for harmonic-complex tones than for single frequency tones. For AM, the harmonic-complex produced a 0.65% signal increase relative to the silent baseline and the single tone produced a 0.47% signal increase. Moreover, for seven of the twelve subjects, the effect of the harmonic-complex tone did reach statistical significance in an individual subject analysis. Therefore, the null result arising from the random-effects group analysis is likely to be due, at least in part, to the high degree of variability in the size and location of the effect observed between subjects.

### *Topographical segregation of AM and FM responses?*

Differential activation by AM relative to FM which was unique to AM occurred in 3 voxels located in a small region of the right posterolateral auditory cortex. We were unable to identify any cortical regions that were specifically FM-dependent. Given that the co-localised activation was by far the dominant pattern of activation, we conclude that the activations by AM and FM broadly overlap. This is at least consistent with the notion that AM and FM stimuli activate common neuronal populations (see Introduction; Mäkelä et al., 1987; Saberi and Hafter, 1995; Eggermont 1994; Gaese and Otswald, 1995; Liang et al., 2002). The fMRI data presented are unable to determine where such a common processing pathway may arise; be it at the cortex or at more peripheral sites within the auditory pathway.

Alternatively, if populations of auditory cortical neurons acted primarily as groups of frequency analysers (filters) tuned to different sinusoidal frequencies, the

low amount of segregation could reflect the spectral similarity of the stimuli. Figure 1(i) illustrates the frequency components of the AM and FM tones. The principal frequency components for both FM and AM tones are centred on the same auditory channel and hence would stimulate common populations of frequency-sensitive auditory cortical neurons.

### *The AM stimulus used in the present study is more potent than the FM stimulus*

Whilst the present study found no clear spatial segregation of areas responsive to either type of modulation, greater auditory activation overall was observed to the AM than the FM stimuli. Moreover, this AM/FM difference was present for both single frequency and harmonic complex carriers. Modulated stimuli were as closely matched as possible for loudness, spectrum and rate and depth of modulation (as described in the Methods section) and so it is unlikely that these cues would underlie the activation differences. The absolute level of the AM stimuli was 1 dB SPL greater than that of the FM stimuli when quantified post-experimentally. This small discrepancy between the overall energies of the two stimuli is unlikely to be sufficient to cause the difference observed between activation to the AM and FM stimuli. The difference in magnitude observed in the present study between the AM and FM harmonic-complex tones was 0.16% signal change. Brechmann et al. (2002) demonstrate, for an area of cortex including the primary area and lateral non-primary cortex, that a 30 dB SPL increase in the sound level of FM tones (from 72 dB to 102 dB) results in a comparable 0.15% signal change. This suggests that an increase in sound level of 1 dB SPL is not sufficient to produce the difference in the measured response.

It is possible that AM is simply a more potent stimulus than FM, and that the posterolateral non-primary auditory cortical regions are more sensitive to changes in amplitude than changes in frequency over time. This is consistent with data presented by Eggermont (1994) from cat primary auditory cortex showing that, for many units, the response was strongest for sinusoidal AM of a carrier tone at the best frequency and weaker, but similar in type, for sinusoidal FM. There is also some psychophysical evidence suggesting that at low modulation rates, where all the frequency components lie within a critical band, FM is less detectable than AM (Zwicker, 1952). Although the AM and FM tones in the present study were both perceived to be clearly modulated.

**Acknowledgements**

**References**

Binder J, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex* 10 : 512-528.

Brechmann, A., Baumgart, F., Scheich, H. (2002) Sound-Level-Dependent Representation of Frequency Modulations in Human Auditory Cortex: A Low-Noise fMRI Study. *Journal of Neurophysiology* 87 : 423–433.

Burkhard MD, Sachs RM (1975) Anthropometric manikin for acoustic research. *Journal of the Acoustical Society of America* 58 : 214-222.

Eggermont JJ (1994) Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hearing Research* 74 : 51-66.

Gaese BH, Otswald J (1995) Temporal coding of amplitude and frequency modulation in the rat auditory cortex. *European Journal of Neuroscience* 7 : 438-450.

Giraud AL,. Lorenzi,C, Asburner J, Wable J, Johnsrude I, Fracowiak R, Kleinschmidt A (2000). Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology* 84 : 1588-1598.

Griffiths TD, Buchel C, Fracowiak RS, Patterson RD (1998). Analysis of temporal structure in sound by the human brain *Nature Neuroscience* 1 : 422-427.

Hall DA, Johnsrude IS, Haggard MP, Palmer AR, Akeroyd MA, Summerfield AQ. (2002) Spectral and temporal processing in human auditory cortex. *Cerebral Cortex* 12 : 140-149.

Hall DA, Haggard MP, Summerfield AQ, Akeroyd MA, Palmer, AR, Bowtell RW. (2001) Functional magnetic resonance imaging measurements of sound-level

encoding in the absence of background scanner noise. *Journal of the Acoustical Society of America* 109 : 1559-1570.

Hall DA, Summerfield AQ, Gonçalves MS, Foster JR., Palmer AR, Bowtell RW. (2000) Time-course of the auditory BOLD response to scanner noise. *Magnetic Resonance in Medicine* 43 : 601-606.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR Gurney E, Bowtell RW. (1999) Sparse Temporal Sampling in auditory fMRI. *Human Brain Mapping* 7 : 213-223.

Harms MP, Melcher JR. (2002) Sound repetition rate in the human auditory pathway : representations in the waveshape and amplitude of fMRI activation. *Journal of Neurophysiology* 88: 1433–1450.

Hart HC, Palmer AR, Hall DA. (2002) Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex. *Hearing Research* 171, 177-190.

Heil P, Rajan R, Irvine DRF. (1992) Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II : Organisation of response properties along the 'isofrequency' dimension. *Hearing Research* 63 : 135-156.

Holmes AP, Friston KJ (1998) Generalisability, random-effects and population inference. *NeuroImage* 7 : S754.

Houtgast T, Steeneken HJM (1985) A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *Journal of the Acoustical Society of America* 77 : 1069-1077.

Kay RH. (1982) Hearing of modulation in sounds. *Physiological Reviews* 62 : 894-975.

Liang L, Lu T, Wang X (2002) Neural Representations of Sinusoidal Amplitude and Frequency Modulations in the Primary Auditory Cortex of Awake Primates. *Journal of Neurophysiology* 87: 2237–2261.

Mäkelä JP, Hari R, Linnankivi A (1987) Different analysis of frequency and amplitude modulations of a continuous tone in the human auditory cortex: a neuromagnetic study. *Hearing Research* 27 : 257-264.

McGonigle DJ, Howseman AM, Athwal BS, Friston KJ, Frackowiak RSJ, Holmes AP. (2000) Variability in fMRI: an examination of intersession differences. *NeuroImage* 11 : 708-734.

Mendelson JR, Schreiner CE, Sutter ML, Grasse KL. (1993) Functional topography of cat primary auditory cortex: Responses to frequency-modulated sweeps. *Experimental Brain Research* 94 : 65-87.

Miller KL, Luh WM, Liu TT, Martinez A, Obata T, Wong EC, Frank LR, Buxton RB (2001) Nonlinear temporal dynamics of the cerebral blood flow response. *Human Brain Mapping* 13 : 1-12.

Møller, AR (1971) Unit responses in the rat cochlear nucleus to tones of rapidly varying frequency and amplitude. *Acta Physiologica Scandinavica* 81: 540-556.

Moore, BCJ, Glasberg BR, Baer, TA (1997) A model for the prediction of thresholds, loudness and partial loudness. *Journal of the Audio Engineering Society* 45 : 224-240.

Moore BCJ, Vickers DA, Baer T, Launer S (1999). Factors  affecting the loudness of modulated sounds *Journal of the Acoustical Society of America* 105 : 2757-2772.

Rees A, Møller AR (1983) Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hearing Research* 10: 301-330

Regan D, Tansley BW (1979) Selective adaptation to frequency-modulated tones: evidence for an information-processing channel selectively sensitive to frequency changes. *Journal of the Acoustical Society of America* 65 : 1249-1257.

Saberi K, Hafter ER (1995) A common neural code for frequency and amplitude-modulated sounds. *Nature* 374 : 537-539.

Thivard L, Belin P, Zilbovicius M, Poline JB, Samson Y (2000) A cortical region sensitive to auditory spectral motion *Neuroreport* 11 : 2969-2972.

Urbas J, Müller-Preuss P, Fastl H, Hesse A, Schorer E (1986) Selective responses of monkey auditory cortex neurons to various parameters of amplitude modulation (AM). *Neuroscience Letters Supplement* 26 : S407.

Wakefield GH, Viemeister NF (1984) Selective adaptation to linear frequency modulated sweeps: Evidence for direction-specific FM channels? *Journal of the Acoustical Society of America* 75: 1588-1592.

Whitfield IC, Evans EF (1965) Responses of auditory cortical neurons to stimuli of changing frequency *Journal of Neurophysiology* 28: 656-672.

Worsley KJ, Friston KJ (1995) Analysis of fMRI data revisited-again, *NeuroImage* 2 : 173-181.

Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. *Cerebral Cortex* 11 : 946-953.

Zwicker, E (1952) Die grenzen der hörbarkeit der amplitudenmodulation und der frequenzmodulation eins tones *Acustica* 3: 125- 133.

**Figure Legends**

**Figure 1. (i)** The left column shows spectrograms of a 1-s section of the six stimuli used. Frequency is represented on the vertical axis with time on the horizontal axis. The sound level of the stimuli, at each point in frequency and time, is represented by the brightness of the greyscale. (A) unmodulated, 300-Hz tone, (B) unmodulated, harmonic-complex tone, (C) frequency-modulated, 300-Hz tone, (D) frequency-modulated, harmonic-complex tone, (E) amplitude-modulated, 300-Hz tone, and (F) amplitude-modulated, harmonic-complex tone. The right column shows the frequency spectra of the same stimuli, with the amplitude of each component on the vertical axis, and frequency on the horizontal axis. **(ii)** Experimental design illustrating the presentation of stimuli (labelled blocks), interspersed with image acquisition (shown by the grey blocks).

**Figure 2.** Black areas indicate auditory activation (P < 0.001) by the different tonal stimuli for three typical subjects, overlaid onto the individual's normalised structural image. The slice orientation is a 34° oblique axial slice, angled along the supratemporal plane and including Heschl's gyrus (HG), whose outline is shown in white on the left-hand image for each subject. The brain slices are displayed in neurological convention (L=L), as illustrated on the top right slice. For an indication of size, a 10 mm scale bar is shown in the bottom right.

**Figure 3.** Auditory activation by amplitude and by frequency modulation. Auditory brain areas, shown in red, are more activated by FM than by unmodulated tones (P<0.001). Areas shown in green were more activated by AM than by

unmodulated tones (P<0.001). Yellow areas are those where the red and green activated regions are overlapping. Outlined in black are those auditory regions which are significantly more activated by AM than by FM tones (P<0.001). Activation was determined using random-effects group analyses and is overlaid onto the standardised mean image for the 12 subjects. The sagittal view (bottom right) illustrates the slice orientation selected for display (axial slice angled by 34º). Slices are displayed in a descending order, viewed at 2.5 mm intervals and are in neurological convention (L=L), as illustrated on the top left-hand slice. A 10 mm scale bar is indicated in the bottom left.

**Figure 4.** Plots of the response magnitude for AM and FM harmonic-complex tones in a region of the left and right auditory cortex. The plots illustrate values for bilateral auditory cortex located on the STP, including HG and immediately adjacent areas. The views are displayed in neurological convention (i.e. the left hemisphere is shown in the left panel) with lateral and posterior boundaries labelled. The orientation of the long axis of Heschl's gyrus is plotted as a red line, where the end points are defined by the 50-75% probability contour for HG (Penhune et al., 1996). **i)** The plots for the AM and FM harmonic-complex tones used in the present study reveal a salient bilateral peak in the auditory region. All peaks are located posterior and lateral to HG, but the elevated response also extends into the lateral part of HG itself. These plots also indicate a greater bilateral posterolateral response to the AM than to the FM tone. **ii)** Data reported by Hall et al. (2001) for a pulsed FM harmonic-complex tone, plotted for the same region of the STP as in i). Again, bilateral peaks of auditory activation in the posterolateral region were found, extending into lateral HG. The

distribution of the response magnitudes are surprisingly consistent, given the different subjects, stimuli and scanning regime.


**Figure 5.** Effect of each tone stimulus on the mean response magnitude within an activated auditory region. This region included all those voxels which displayed significant effects of both amplitude and frequency modulation, relative to the unmodulated tones and corresponds to the yellow area shown in Figure 3. The mean percentage signal change was calculated for each tone stimulus relative to the silent baseline, for all subjects. Open circles denote the single-frequency tones and filled circles denote the harmonic-complex tones. Error bars denote the 95% confidence limits for the mean.
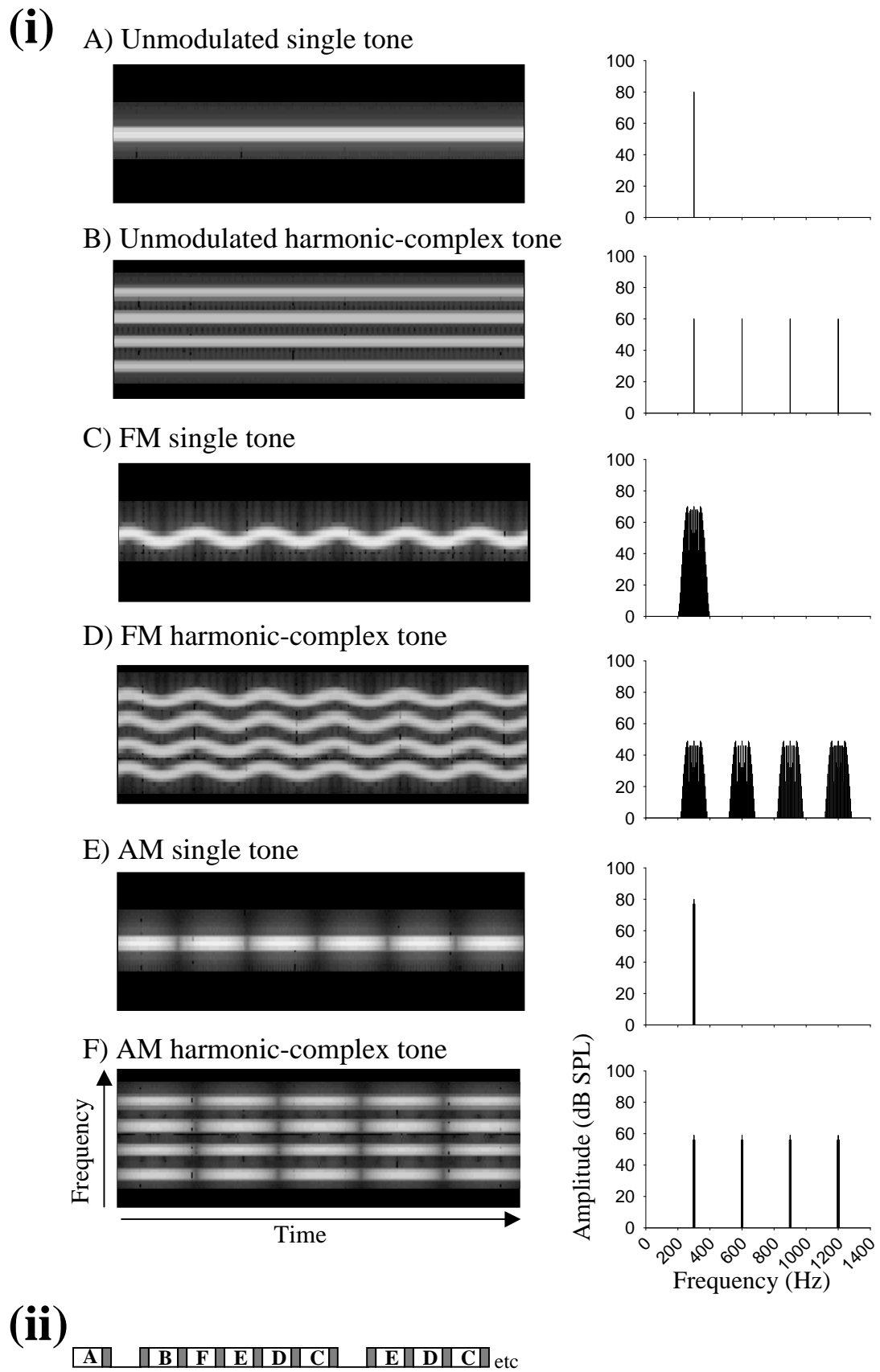
**(i)**
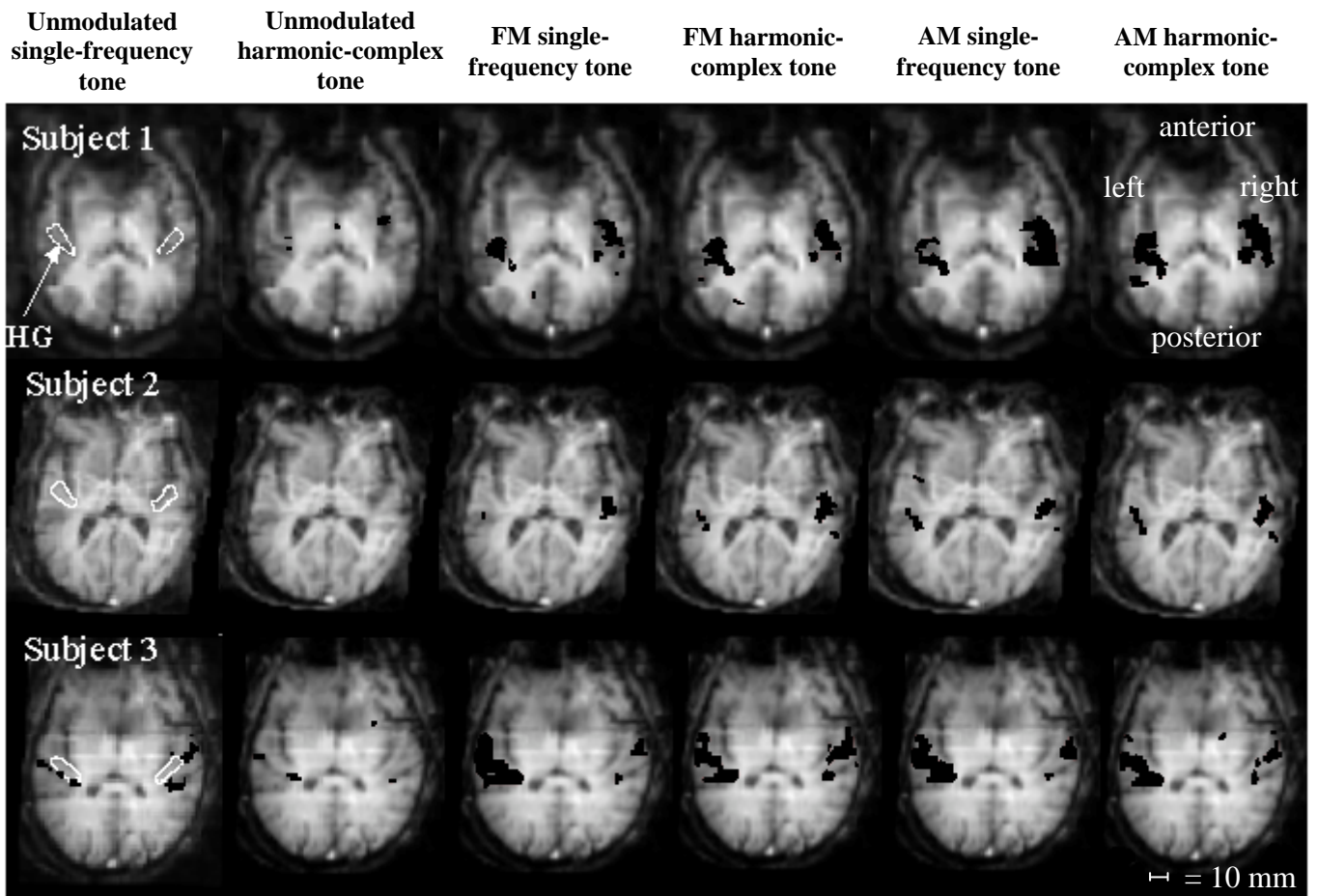
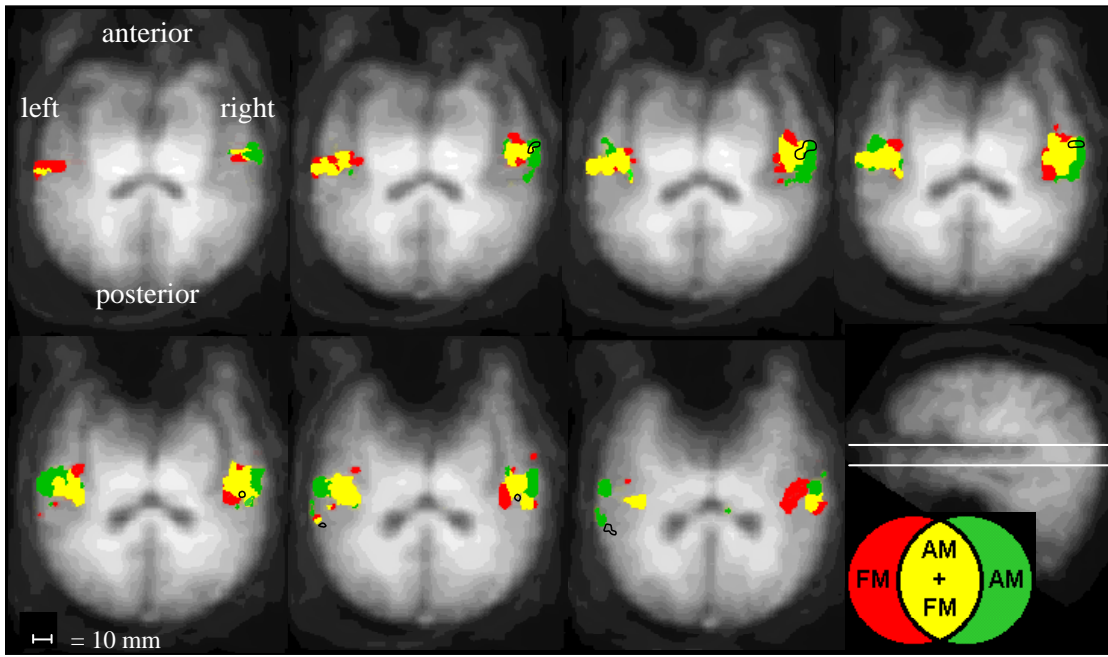A) Unmodulated single tone



B) Unmodulated harmonic-complex tone



C) FM single tone



D) FM harmonic-complex tone



E) AM single tone



F) AM harmonic-complex tone



Frequency

Time

Amplitude (dB SPL)

Frequency (Hz)

**(ii)**

| A | | B | F | E | D | C | | E | D | C | etc |

Figure 1

| Unmodulated single-frequency tone | Unmodulated harmonic-complex tone | FM single-frequency tone | FM harmonic-complex tone | AM single-frequency tone | AM harmonic-complex tone |
|---|---|---|---|---|---|



Figure 2

Figure 3

**(i)**

AM harmonic-complex tone

FM harmonic-complex tone

**(ii)**

pulsed FM harmonic-complex tone
(Hall et al., 2001)

Standardised
response
magnitude

-1
0
1
2

Figure 4.

Figure 5