Asymmetrical neural processing of amplitude modulated sounds: a psychophysical, fMRI and TMS investigation.

Adam Michael Partridge BSc, MSc

PhD

University of York

Psychology

September 2018

Abstract

Human speech is the most behaviourally important and complex signal that the human brain is required to process yet it does so with remarkable ease. Speech is composed of highly complex amplitude modulations over time and these modulations are known to be crucial for intelligibility. There is evidence for hemispheric asymmetries in processing auditory modulations over different timescales and multiple models have been proposed to account for these. The procedure by which the auditory system extracts and processes these modulations is not fully understood. Psychophysical, neuroimaging and non-invasive neurostimulation techniques can be combined in complementary ways to potentially provide unique insights into this problem. Functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) are relatively novel methods that have not been previously applied in combination to investigate amplitude modulation processing. Three psychophysical and fMRI-guided TMS studies were conducted in order to address the following research questions. Firstly, is fMRI-guided TMS an effective method for modulating AM processing? Secondly, are different TMS protocols more or less effective at modulating AM processing? Finally, is fMRI-guided TMS an effective method for further understanding the functional asymmetry of speech processing? Online dual pulse TMS to right auditory cortex was shown to be effective at modulating 4 Hz AM detection accuracy. State-dependent TMS to left auditory cortex was shown to be effective at modulating 40 Hz AM detection accuracy, but the effects were complex. Continuous theta burst stimulation was not shown to be effective at modulating AM depth discrimination ability. It was thus found that fMRI-guided TMS can be an effective tool for modulating AM processing, however, efficacy differs depending on the specific TMS protocol used. Further, fMRI-guided TMS can be used to investigate functional asymmetry of speech processing, however some important caveats apply.

List of Contents

List of Tables

List of Figures

Author's Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 4 has been published as a pre-print on the Open Science Framework and has been submitted to a peer-reviewed journal for publication:

Partridge, A. M., Prendergast, G., & Hymers, M. (2018, April 6). Amplitude Modulation detection ability can be differentially affected using state-dependent Transcranial Magnetic Stimulation. https://doi.org/10.31219/osf.io/twfc5

Chapter 1

Introduction

1.1 Overview

The human ability to communicate using speech is of vital importance. The fact that information can be encoded in minute vibrations of the air and that these vibrations can be decoded with such little conscious effort from the listener is remarkable. This ability is yet more remarkable when the complexity of encoded information, the transmission speed and the fine motor skill required to produce speech is considered. Processing natural speech is a hugely complex process and disorders of auditory processing often lead to profound communication difficulties that significantly affect quality of life. Despite the importance and ubiquity of speech, mechanisms underlying the processing of its components are still not fully understood.

The acoustic speech signal is comprised of complex, continuous and concurrent modulations of amplitude and frequency. The processing of these low-level aspects of speech is of particular interest as all higher-order linguistic complexity must be represented by these, relatively few, parameters of interest. In isolation, neither temporal nor spectral modulations can form an intelligible speech percept (e.g. Shannon et al. 1995), but determining the relative contribution of these features is an important goal.

Amplitude modulations (AM) of the speech envelope have been shown to be crucial for speech intelligibility (Drullman et al. 1994a,b). Processing of AM in the subcortical stages of the auditory system is well-characterised (e.g. Frisina 2001) however multiple accounts have been proposed to explain how this low-level aspect of speech is processed by the brain. Theories of cortical AM processing must account for the observed structural and functional asymmetries in auditory cortical regions during auditory processing. The application of non-invasive neuroimaging and neurostimulation techniques, such as functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) to investigate behaviour is a relatively recent development and it is not currently known how effective these techniques are for modulating sensitivity to amplitude modulation. Using non-invasive brain imaging and stimulation methods to investigate amplitude modulation processing could therefore provide unique and novel insights. The work presented here aims to further investigate the processing of amplitude modulation through the combination of behavioural, neuroimaging and brain stimulation techniques.

This investigation was designed to address a set of three overarching research questions.

- Firstly, is fMRI-guided TMS an effective method for modulating AM processing?
- Secondly, does the effectiveness of TMS for modulating AM processing differ based on the TMS protocol used?
- Finally, is fMRI-guided TMS an effective method for further understanding the functional asymmetry of speech processing?

1.2 The Acoustic Speech Signal

Speech is a highly complex signal comprising of continuous and concurrent modulations of amplitude and frequency.

The lower section of Figure 1.1 shows how the frequency and amplitude of a

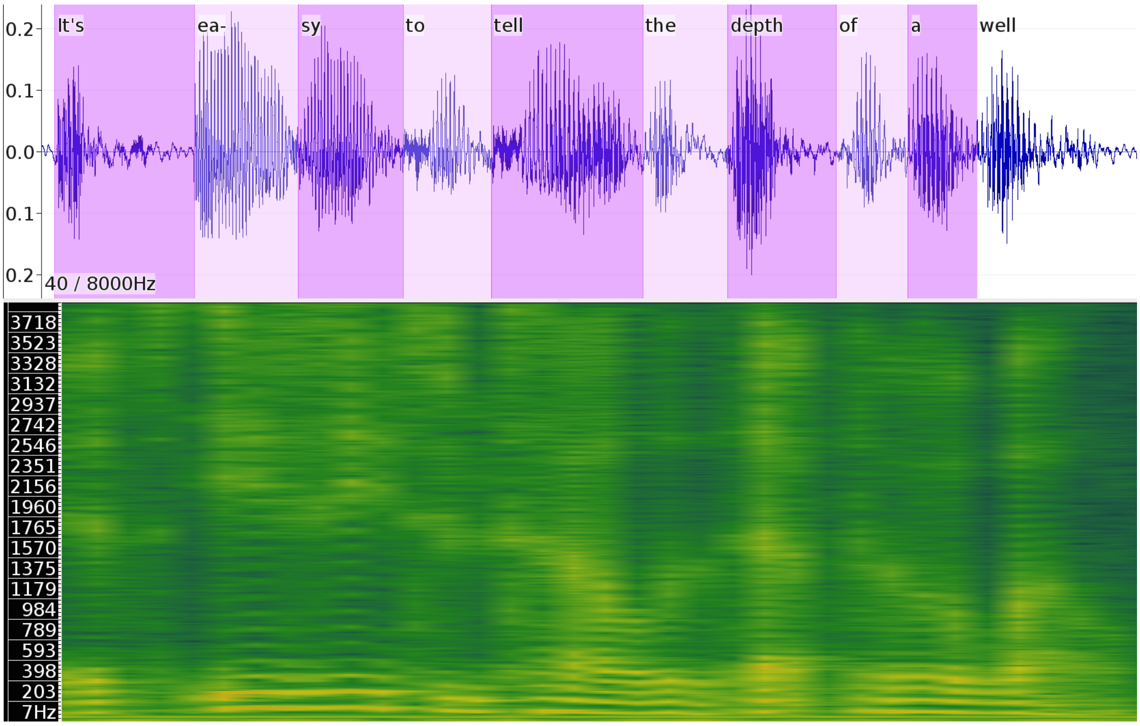

Figure 1.1: Upper: waveform representation of speech with amplitude on the Y axis. Lower: spectrogram representation of speech with frequency on the Y axis. Time is on the shared X axis. The wave file for this sentence was sourced from the *Open Speech Repository*.

speech signal changes over time, with time and frequency components visible as bands of pulsing colour (in horizontal and vertical directions, respectively). The upper section of Figure 1.1 shows how the amplitude of the speech signal changes over time. The rate of signal fluctuations in this time domain has been used to categorise temporal aspects of speech (Rosen 1992). The speech signal has been dissociated into envelope, periodicity, and temporal fine structure components. The envelope is the lowest frequency component and refers to overall fluctuations in amplitude at $\sim 2 - 50$ Hz. The envelope contains a substantial amount of useful speech information, for example, prosodic and manner of articulation cues. Specific prosodic information relating to stress and intonation is contained in the periodicity information. This component is subdivided again into periodic and aperiodic stimulation. Periodic sounds fluctuate at frequencies between $\sim$50 and 500 Hz whereas aperiodic sounds have a less well defined frequency range, from a few kHz upwards. Temporal fine structure contains fluctuations in the range ~600 Hz – 10 kHz and is responsible for place of articulation information, vowel quality and segmental cues to voicing and manner. These subdivisions are useful when discussing aspects of temporal modulations in speech but it is important to make a distinction between the low frequency information extracted analytically and the low frequency envelope of speech. An analytic extraction (e.g. with a Hilbert transform) will typically combine envelope and periodicity information, so many accounts disregard periodicity as a separate component (e.g. Joris et al. 2004; Moon and Hong 2014). Figure 1.2 shows a full speech signal and its envelope and fine structure components.

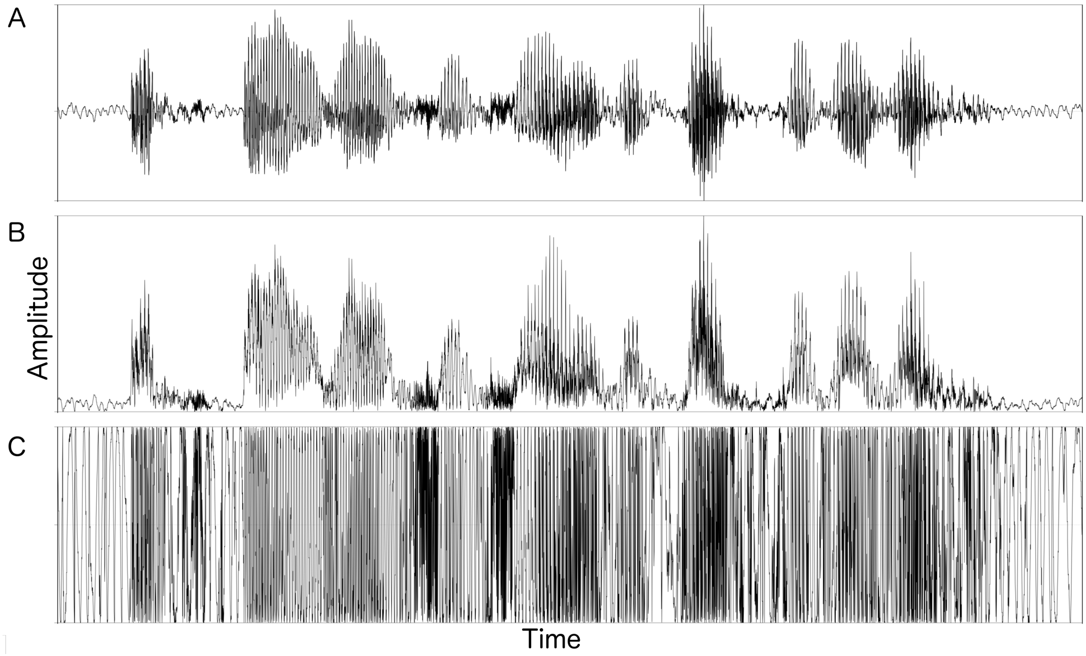

Figure 1.2: Hilbert extraction of envelope (B) and temporal fine structure (C) from a speech signal (A). The sentence is the same as seen in Fig 1.1.

In order to investigate the relative role of these features and how they contribute to intelligibility, one informative behavioural approach involves isolating parameters of interest and then systematically modifying them and measuring subsequent intelligibility. By measuring the marginal contribution of these parameters, it is possible to ascertain which are most critical to intelligibility and therefore transmit the most useful information. In order to investigate the relative importance of amplitude and frequency information, intelligibility has been measured while the spectral content of speech was severely degraded. Using band-pass filters to limit the amount of spectral information, and low-pass filters to extract the low-frequency information from these bands, it was found that despite a reduction in spectral content to just 3 bands, 90% intelligibility could be retained, provided the full speech envelope from each band was reimposed on it before recombination (Shannon et al. 1995). Other estimates for the number of bands required for intelligible speech are similar, e.g., asymptotic performance at 5 – 8 bands, depending on stimulus difficulty (Dorman et al. 1997); 90% intelligibility with 5 channels, asymptotic performance at 8 bands (Loizou et al. 1999). A similar approach involved measuring the speech reception threshold for speech signals low- and high-pass filtered at different cut-off frequencies. AM rates below 16 Hz were found to be the most crucial for intelligibility (Drullman et al. 1994a,b). Elliott and Theunissen (2009) found that AM at 1 – 7 Hz was most critical for speech intelligibility and it has since been shown that cross-channel AM sweeps may also be important (Prendergast and Green 2012). While AM of limited frequency content may be sufficient for intelligibility in quiet, the additional frequency information may be required or utilised more for challenging tasks such as speaker gender identification or speech recognition in background noise (Elliott and Theunissen 2009; Moon and Hong 2014).

To summarise, amplitude modulations are present in speech at a wide range of rates, but not all rates are of equal importance for the intelligibility of a speech signal. These behavioural studies allude to the large amount of redundancy in the speech signal but primarily demonstrate that low frequency AM of the speech envelope is crucial for speech intelligibility.

1.3 Controlled Auditory Stimuli

An alternative to using natural or filtered speech stimuli is using controlled auditory signals. This approach allows the isolation of key parameters of interest and these controlled stimuli, such as amplitude and frequency modulated signals, serve as useful intermediates for developing models of speech processing. Using this approach, individual low-level aspects of speech can be isolated and systematically varied in studies, in order to investigate the marginal effect of that single parameter on a behavioural or physiological measure. The signals can be adjusted to approximate specific perceptual properties of speech and they can be studied in isolation or combination, e.g. to model concurrent spectro-temporal processing (Chi et al. 1999; Luo et al. 2006; Hsieh and Saberi 2010; Edwards and Chang 2013). An amplitude modulated signal is described by Equation (1.1).

$$signal = A_0 * (1 + mod) * carrier \tag{1.1}$$

A_0 is a scaling factor that controls the amplitude of the signal. *mod* is the modulator signal, which determines the overall envelope of the amplitude modulation. *carrier* is the carrier signal, which determines the signal to be modulated.

Sinusoidal amplitude modulation (SAM) is commonly used to experimentally model amplitude modulations in speech (Joris et al. 2004). The use of SAM is physiologically supported as the modulations present in the speech spectrogram are characterised by near-sinusoidal patterns (Singh and Theunissen 2003; Elliott and Theunissen 2009). A further rationale for using this signal is that the response of linear systems to complex input can be inferred by measuring the system response to one SAM stimulus and then progressively adding more. This rationale is somewhat limited by non-linear processing stages of the auditory system. For example, evidence for non-linear neural encoding of concurrent amplitude and frequency modulation has been found, and computational modelling results suggest this may reflect a tradeoff between encoding these features (Luo et al. 2006; Santoro et al. 2014). However, when patterns of linear processing are found, such as AM processing as reflected by the auditory steady state response (e.g. Prendergast et al. 2010), theis approach can be effective. Sinusoidal amplitude modulation (SAM) is described by Formula Group (1.2).

$$carrier(t) = sin(2\pi t f_c) \tag{1.2a}$$

$$mod(t) = sin(2\pi t f_m) \tag{1.2b}$$

$$signal = A_0 * (1 + mod) * carrier \tag{1.2c}$$

$$x(t) = A_0(1 + mod * sin(2\pi t f_m)) * sin(2\pi t f_c) \tag{1.2d}$$

The carrier signal determines the fine structure of the resulting output signal; the carrier type in Formula Group (1.2) is sinusoidal, resulting in a SAM tone with a centre frequency of f_c. Modulation rate refers to the frequency of the AM. Suprathreshold SAM at frequencies <~10 Hz is perceived as a slow fluctuation in amplitude, whereas SAM ~10 – ~50 Hz has a perceptual 'unevenness' or 'roughness' quality and SAM >~50 – 100 Hz is perceived as tonal with an increase in pitch with further rate increments (Zwicker 1952). Edwards and Chang (2013) thoroughly review the literature describing how AM perception changes as rate (and carrier frequency) increases.

The depth of amplitude modulation can be quantified as the modulation index m. Modulation index m is described by Equation (1.3).

$$m = \frac{(\text{peak}A_m)}{(\text{peak}A_c)} \tag{1.3}$$

peakA_m is the peak amplitude of the modulator signal and peakA_c is the peak amplitude of the carrier signal. Figure 1.3 shows three SAM noise carriers modulated at different depths. Modulation depth is often reported in units of 20log$m$

dB. For convenience, some converted values are given here: -30dB = $0.032m$; -20dB = $0.1m$; -10dB = $0.316m$; and 0dB = $1m$. As the depth of amplitude modulation decreases, it becomes more difficult to detect. The modulation depth of an AM signal can be systematically varied in order to assess sensitivity to AM.

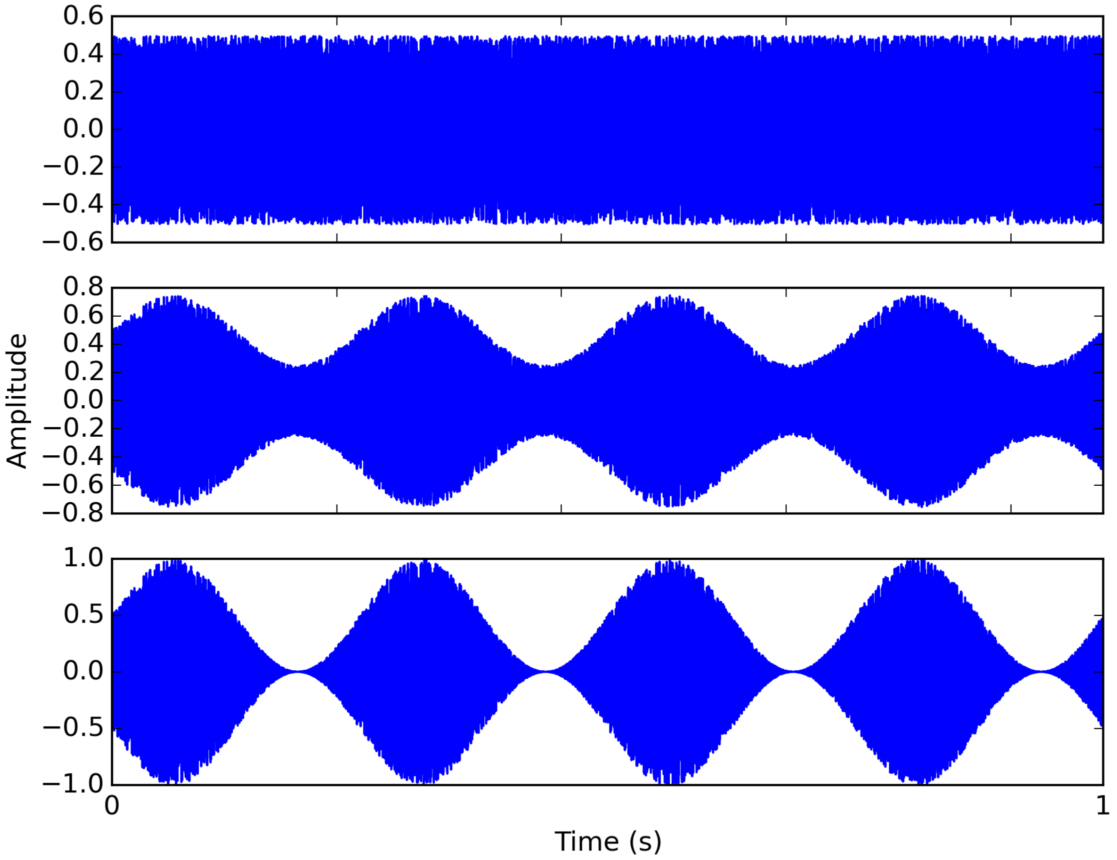

Figure 1.3: Noise carriers sinusoidally amplitude modulated at 4 Hz, with different m values. Upper: $m = 0$. Middle: $m = 0.5$. Lower: $m = 1$.

To summarise, controlled auditory stimuli are often used to model low-level aspects of speech. This approach allows fine experimental control over variables of interest. SAM is most commonly used to approximate the amplitude modulations present in speech. The modulation index is a key parameter which represents the depth of modulation, with lower depths more difficult to detect. Systematically modifying this parameter allows researchers to measure sensitivity to AM.

1.4 Behavioural Investigations into the Detection of Amplitude Modulation

The earliest studies that investigated AM of the temporal envelope used the phenomenon of perceptual ‘beats’ to create the modulations. This is where two concurrent sinusoidal tones that differ in frequency are perceived as a singular tone that fluctuates in amplitude. The rate of the perceived AM is equal to the frequency difference between the tones. For example, if a 495 Hz and 505 Hz tone are played concurrently, a 500 Hz tone would be perceived, with an amplitude fluctuation at a rate of 10 Hz (Licklider et al. 1950).

Riesz (1928) used this phenomenon to investigate the sensitivity of the auditory system and found that it was most sensitive to AM changes at a rate of $\sim$ 3Hz. However, this method is disadvantaged by a limited carrier frequency bandwidth and at modulation rates greater than 30 Hz, two distinct tones are perceived, limiting the AM rates at which the technique can be used (Licklider et al. 1950). Technological advances allowed later researchers to directly modulate more complex signals, as described in Equation (1.1). Another early study used this technique to measure the lowest level of SAM that could be detected by the auditory system and found a similar peak sensitivity at 4 Hz (Zwicker 1952). This account is the most complete single study of AM detection thresholds to date, with a wide range of AM rates (1 – 6000 Hz) and three carrier frequencies (0.25, 1, and 4 kHz). These findings provided early indications of the importance of low frequency modulations in the auditory domain.

The temporal modulation transfer function (TMTF), describes the relationship between AM detection threshold and AM rate (Rodenburg 1977; Viemeister 1979). TMTFs have been measured using both tone (e.g. Kohlrausch et al. 2000) and noise carriers (e.g. Viemeister 1979) and both measures have benefits for understanding the processes involved in speech perception. Briefly, sinusoidal carriers allow investigation of the effect of carrier frequency, which may elicit information about secondary filtering processes, however, it is difficult to control which cues participants use to detect the modulation (e.g., at high modulation frequencies where the sidebands are resolved on the cochlear membrane and the perception of modulation is diminished). Noise carriers do not suffer from this problem however, the intrinsic random fluctuations of these signals complicates interpretation, particularly with narrowband noise signals. This distinction allows researchers to specifically explore the relative role of internal and external limitations (arising from the stochastic nature of the noise signals) on the processing of AM (Ewert and Dau 2004). In order to fully understand the perception of amplitude modulation by the auditory system, it is necessary for studies with both tone and noise carriers to be performed (Kohlrausch et al. 2000; Joris et al. 2004).

1.4.1 TMTFs measured with noise carriers

Rodenburg (1977) measured the TMTF using noise carriers with AM rates from $\sim 3 - 1000$ Hz and noted that sensitivity was high and constant below 10 Hz but slowly decreased above that rate. He concluded that the TMTF was low-pass in shape with a cutoff at $\sim$ 50 Hz. A later study also measured the TMTF using a wideband noise carrier and modulation rates from 2 – 4000 Hz and found evidence to support this account (Viemeister 1979). This study also suggested the TMTF was low-pass in nature (though a subtle decline in sensitivity was shown for rates below 4 Hz), with a -3dB cutoff at $\sim$ 60 Hz. These studies also found that overall intensity had no effect on the response. A later study measured AM detection for octave-wide noise bands at f_m values from 1 – 64 Hz and found a similar characteristic in the TMTF, with peak sensitivity at 1 – 4 Hz (Ozimek and Sek 1988). Gutschalk et al. (2008) measured AM detection of broadband noise modulated at rates from 50 – 500 Hz and provided further support for the low-pass account, although the sensitivity to low rates (speech envelope rates and below) were not measured. Figure 1.4 shows TMTFs estimated with broadband noise carriers. Overall, these studies show a peak sensitivity to low rates of AM ($<\sim$10 Hz) and sensitivity to AM attenuates as modulation rate increases.

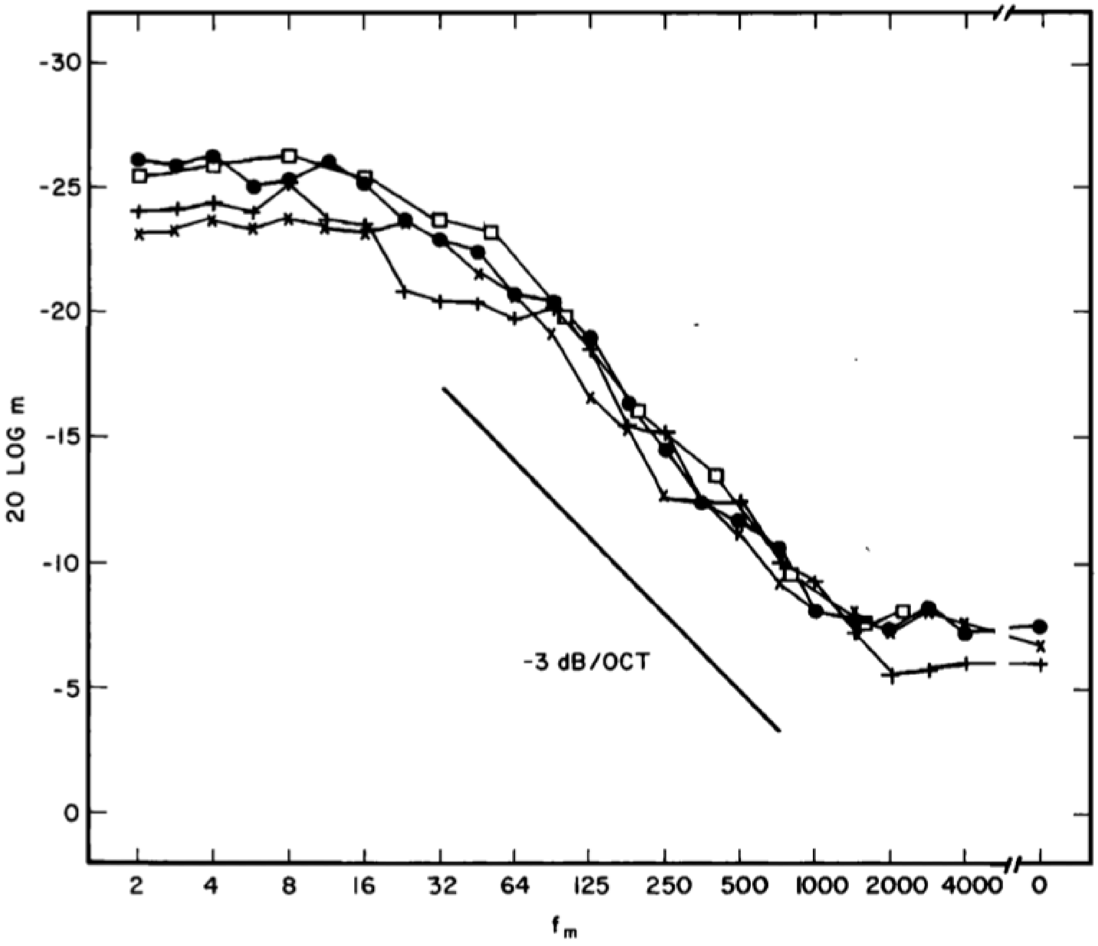

Figure 1.4: Temporal modulation transfer functions measured with noise carriers. Modulation rate in Hz is represented on the X axis. Modulation depth in 20log$m$ dB is represented on the Y axis. Reproduced with permission from Viemeister (1979). Copyright 1979, Acoustical Society of America.

1.4.2 TMTFs measured with tone carriers

A comprehensive study of AM detection measured the TMTF with sinusoidal carriers from 1 – 10 kHz and AM rates from 10 – 1600 Hz (Kohlrausch et al. 2000). This study measured a relatively flat threshold for AM rates up to ∼130 Hz, followed by a steady (approximately 6dB/octave) increase in threshold up to ∼800 Hz. However, when each carrier reached its critical frequency (Schorer 1986; Sek and Moore 1994), participants could detect the spectral sidebands and sensitivity increased again. This critical modulation frequency differed for each carrier frequency; see Figure 1.5. In contrast to TMTFs measured with noise carriers, they also found that thresholds decreased with increasing carrier sound level (from 20 – 75 dB SPL). Moore and Glasberg (2001) measured AM detection thresholds using sinusoidal carriers in healthy participants using two carrier levels (30 and 80 dB SPL) and three carrier frequencies (1000, 2000 and 5000 Hz). Their TMTFs resembled those found in earlier studies with relatively flat thresholds until the sidebands were resolved and that sensitivity was greatest in the 80 dB SPL condition. They conducted the same experiment with hearing impaired (HI) listeners and found that HI participants had similar thresholds at 80 dB SPL but actually showed finer sensitivity than healthy controls in the quieter condition. Their results suggest that HI listeners may not be able to perceive the spectral sidebands as well as normal controls at higher AM rates, but overall temporal resolution abilities between HI and normal hearing participants are similar. This somewhat paradoxical finding was also shown more recently when the variability of AM detection within a clinically normal hearing population was investigated. AM sensitivity was shown to vary in a noise-exposed group that had near-normal hearing thresholds, and this group showed consistently lower AM detection thresholds (Stone and Moore 2014). Even within the normal hearing population, variation in temporal processing skills exists and can be linked to life experiences such as occupational noise exposure (Stone et al. 2008; Kumar et al. 2012). Understanding this 'hidden hearing loss,' is the topic of much current research as it has implications for real-world auditory perception (e.g. speech processing in noise) that standard audiometric measures do not adequately capture (Guest et al. 2018; Dewey et al. 2018).

1.4.3 Current views on TMTFs

Interest has largely shifted away from studies directly characterising the TMTF, however, it has been recently revisited due to renewed interest in the interaction of AM sensitivity and syllabic speech rates (Edwards and Chang 2013). Properties of the AM detection TMTF are generally accounted for and this low-pass (i.e. leaky integrator in the time domain) view is commonly held. The low-pass

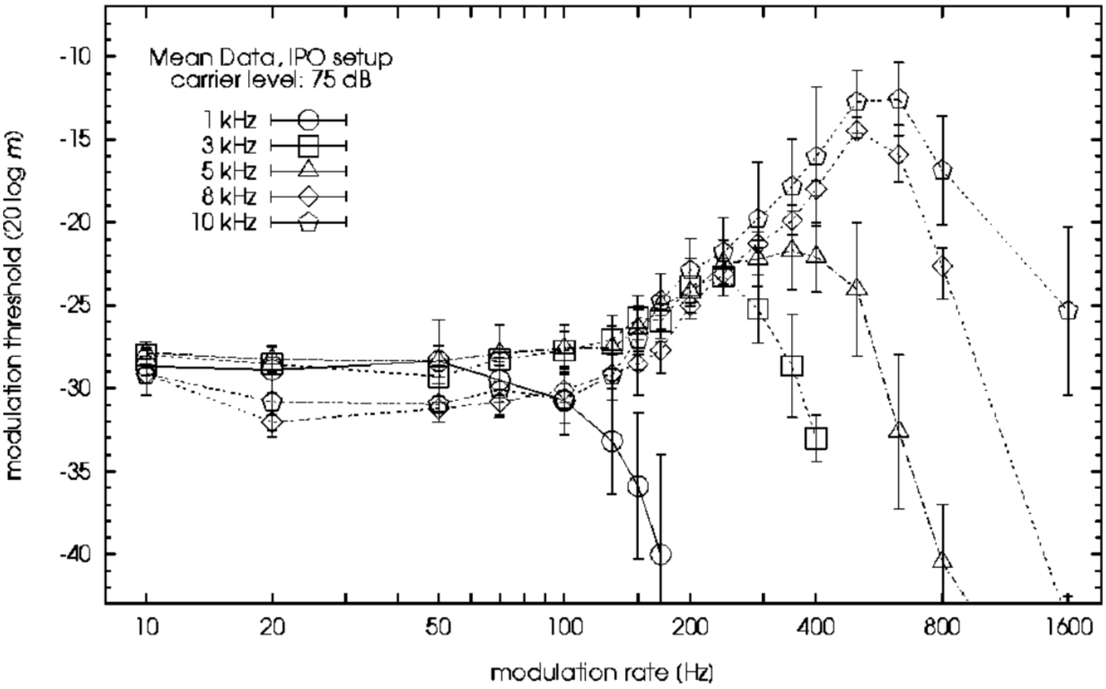

Figure 1.5: Temporal modulation transfer functions for different tone carrier frequencies. The sharp increase in sensitivity represents the critical frequency at that carrier frequency. The y-axis units are dB: -30 dB = 0.032 m. Reproduced with permission from Kohlrausch et al. (2000). Copyright 2000, Acoustical Society of America.

account generally persists despite the findings of a recent review that also found a high-pass component for very low frequencies; forming a band-pass filter with a bandwidth from $\sim 2 - 5$ Hz (Edwards and Chang 2013). This band-pass proposal does not contradict the early work, as few studies examined these very low AM rates, so the function largely showed a low-pass characteristic by these measures. The $2 - 5$ Hz pass band is notable as it corresponds to the syllabic rate of speech (e.g. Houtgast and Steeneken 1985; Krause and Braida 2004), which is remarkably consistent across languages (Ding et al. 2017), suggesting the existence of an underlying universal principle for the optimisation of information transfer with speech.

To summarise, AM detection tasks aim to measure the smallest AM depth of a tone or noise stimulus that can be detected. When applied across multiple AM rates, it is possible to measure the TMTF, with peak sensitivity at roughly 4 Hz, corresponding closely to the syllabic rate of speech (Krause and Braida 2004). It is generally accepted that the TMTF is low-pass in shape, however this may be due to a lack of data for very low rates where a decrease in sensitivity has been found, suggesting a band-pass characterisation may be more accurate. Hearing impairment affects temporal processing non-linearly and individuals with clinically normal hearing thresholds can show temporal processing deficits as measured by AM detection tasks.

1.5 Behavioural Investigations into the Discrimination of Amplitude Modulation Depth

The ability to discriminate between AM depths provides a lower bound at which information can be decoded from the speech signal (Wakefield and Viemeister 1990). The AM discrimination threshold reflects the smallest change in modulation depth that can be detected between two AM signals (i.e., the just-noticeable-difference). In contrast, the AM detection threshold only provides information on the *detection* of the presence of amplitude modulations; it does not provide any detail on how much useful information can be extracted. Wakefield and Viemeister (1990) give an example where AM detection is successful but no information about the depth is preserved. In this case, information transfer is poor and (e.g., syllabic) discrimination errors may be made. Thus, the ability to discriminate AM depth limits the successful transfer of auditory information, such as when humans communicate using speech. In AM depth discrimination studies, periods of AM sounds are compared, e.g. in 2- or 3-alternative forced choice designs (e.g. Grantham and Bacon 1988). One period contains AM at a standard depth and the AM depth of the other period varies. The task is to determine which of the periods contains modulation at a greater depth (i.e. has a larger modulation index, m). The standard depth remains constant within each task and AM depth discrimination thresholds can be plotted as a function of standard depth. This provides a measure of how sensitive the auditory system is to differences in AM depth across the range of detectable depths. This can inform researchers about the relative importance of different rates of AM to the auditory system and can lead to new models and theories that further collective understanding of speech processing.

The caveats surrounding choice of carrier type apply equally to AM depth discrimination tasks and studies have used sinusoidal, narrowband noise and broadband noise carriers. Fleischer (1980) used sinusoidal tone carriers and obtained an AM depth discrimination function for standard depths from -14 – 0 dB 20log$m$. A comprehensive study used octave wide noise carriers (centre f_c 250 – 8000 Hz) and obtained a function from -12 – 0 dB for f_m values of 1 – 64 Hz (Ozimek and Sek 1988). Using broadband noise carriers (0 – 8000 Hz) another function was measured using f_m values of 25, 50 and 400 Hz (Wakefield and Viemeister 1990). In a similar range of standard depths as the earlier studies, comparable results were found. The results for these three studies can be seen in Figure 1.6; a decrease in threshold with a decrease in standard depth. Lee and Bacon (1997) varied the stimulus lengths and at comparable lengths to those used in previous work, a similar function shape was obtained. A more recent study sought to consolidate this evidence using sinusoidal and noise carriers and examine the relative effect of internal and external limitations on AM processing (Ewert and Dau 2004). In their first experiment, AM depth discrimination thresholds were obtained using standard depths from -28 – -3 dB. Their findings supported the established AM depth discrimination functions and highlighted the subtle differences between sine and noise carriers as the standard depth approached detection threshold. In summary, for both sinusoidal and noise carriers, a decrease in threshold was observed as the standard depth decreased. This decrease was a constant fraction of the AM depth down to -18 dB and -8 dB for the sine and noise carrier respectively, following the Weber fraction. However, past this point, a fixed increase in AM depth was required that was independent of the standard depth. Various models have been proposed to account for this pattern, such as the leaky-integrator model, the envelope spectrum model and the performance model (Wakefield and Viemeister 1990; Ewert and Dau 2004).

To summarise, AM depth discrimination tasks aim to quantify the smallest discernible difference in AM depth. When repeated across a range of standard depths, a modulation transfer function can be measured. The findings from AM depth discrimination studies concur with AM detection studies and provide further support for the auditory system's particular sensitivity to low rates of AM.

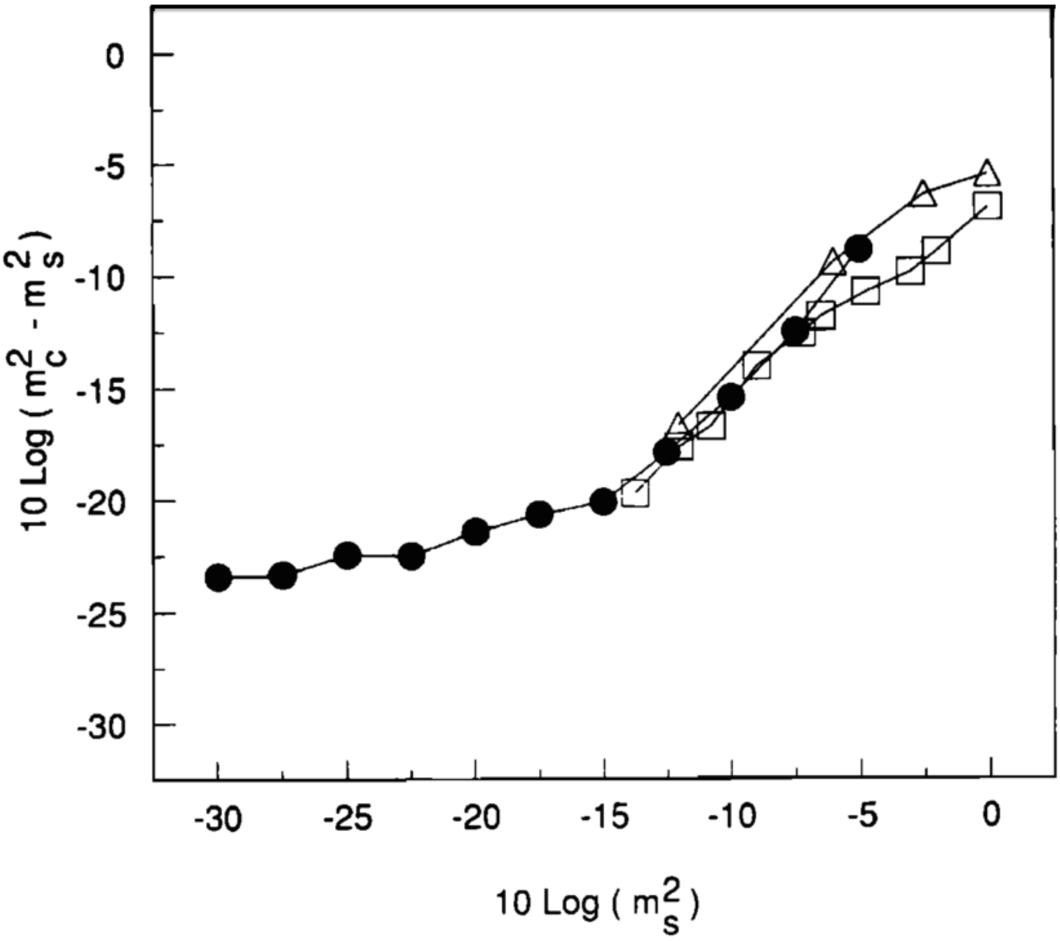

Figure 1.6: AM depth discrimination thresholds as function of standard depth. Circles: Wakefield and Viemeister (1990), f_m = 25 Hz, spectrum level of 25 dB. Squares: Fleischer (1980), f_c = 1000 Hz sinusoid. Triangles: Ozimek and Sek (1988), average of four carrier frequencies. Reproduced with permission from Wakefield and Viemeister (1990). Copyright 1990, Acoustical Society of America.

1.6 Effects of Selective Adaptation on the Perception of Amplitude Modulation

Studies have shown that previous exposure to a stimulus reduces the ability to detect the same stimulus on subsequent trials, known as selective adaptation (Hood 1972). This technique, described as the “psychophysicist’s microelectrode,” (Frisby 1979), allows researchers to systematically probe neural systems using purely behavioural measures. By comparing the adaptation effect of a stimulus to that of a stimulus that varies systematically from the adaptor, for example, comparing AM with frequency modulation (Kay and Matthews 1972; Regan and Tansley 1979) or with AM at a different rate (Wojtczak and Viemeister 2003), it is possible to dissociate neural systems (Hood 1972; Grill-Spector et al. 2006). This method has been used to show that previous exposure to AM tones leads to subsequent decreases in AM sensitivity. This decrease equates to a threefold increase in threshold from baseline for rates from $3 - 30$ Hz, provided that the adapter and test stimulus were modulated at the same rate (Kay and Matthews 1972). The rates of the test stimuli were modified and the effect attenuated as the difference between the adapter and test stimulus modulation rate increased. Subsequent research has supported this account, but has shown that the time needed to reach maximum adaptation is much longer; $600 - 1200$ seconds, compared to 12 seconds (Regan and Tansley 1979). The effect has been examined in more detail and was shown to be most prominent at 16 Hz and no effect of carrier level was found using 30, 50 and 80 dB SPL (Tansley and Suffield 1983). More evidence has been found for a modulation rate sensitivity in a study that examined the effect of adaptation on suprathreshold AM (Wojtczak and Viemeister 2003). More detailed parameters such as the time course of adaptation release, the specificity of rate selectivity and the effect of combined AM are now well characterised (Wojtczak and Viemeister 2005). It has also been found that the effect can be modulated by attention, suggesting the involvement of a top-down component (Kawashima 2009). Adaptation to AM was also found to be induced by pulse trains and this was shown to have an adverse effect on AM detection thresholds at a high modulation rate of 100 Hz. Unlike the low frequency (<∼30 Hz) adaptation to AM described previously (e.g. Tansley and Suffield 1983) this effect did not transfer interaurally, suggesting the involvement of a distinct mechanism from a different stage of auditory processing (Gutschalk et al. 2008). The mechanisms underlying the behavioural effect of adaptation to AM are still not fully understood (Kleinschmidt and Jaeger 2016) however, this approach appears to support the existence of channels that are selective for processing specific rates of modulation. This modulation rate sensitivity led to the proposition of a modulation filterbank stage in models of amplitude modulation processing (Dau et al. 1997a; Ewert and Dau 2000).

To summarise, the phenomenon of selective adaptation can be exploited to investigate neural systems with purely behavioural experiments. This method has been successfully applied to investigate AM processing and there is evidence for specific channels tuned to AM rate in the auditory system.

1.7 Subcortical Processing of Amplitude Modulation

A detailed understanding of how sound is represented in the peripheral and subcortical auditory system, and how AM modifies this representation, is useful as the information the auditory system preserves informs researchers about the relative importance of disparate aspects of the acoustic speech signal. Further, the output from early peripheral stages serves as input to later cortical stages so robust models allow researchers to develop ever finer investigations into the stage of interest.

1.7.1 Peripheral Auditory System

The peripheral auditory system begins at the eardrum and ends at the auditory nerve. In brief, the auditory periphery is responsible for the mechanotransduction of the vibrations that constitute the speech signal into a series of electrical pulses (Young 2008; Verhulst et al. 2018). This early representation is largely faithful to the input signal by directly representing complex modulations (Wang et al. 2003). The cochlea analyses sounds in the frequency domain and initiates a tonotopic organisation that persists throughout the auditory system (Saenz and Langers 2014). Within the cochlea, high frequencies resonate at the base and progressively lower frequencies resonate towards the apex (Gold and Pumphrey 1948). Analytically, signals can be represented in the frequency domain through a Fourier transform, which calculates the power and frequency of each sinusoidal component. This analytical transformation is an adequate model of the cochlea for understanding how sound and AM is initially coded, although more complex models have been developed (Jepsen et al. 2008; Verhulst et al. 2018). In the frequency domain, a pure tone manifests as a single peak at the carrier frequency, f_c, because all of the energy in the signal is found at this frequency. Figure 1.7 shows the frequency spectrum of three SAM pure tone signals; each signal is clearly the sum of three distinct frequency components. The spectrum takes the form of one dominant peak at the carrier centre frequency, f_c, and two spectral sidebands with less power. The lower sideband manifests at a frequency of $f_c - f_m$ and the upper sideband manifests at a frequency of $f_c + f_m$. For example, a 500 Hz sinusoidal carrier with SAM at 10 Hz would have a main peak at 500 Hz and two sidebands at 490 Hz and 510 Hz. The power of the sidebands are equal, and their power relative to the f_c peak is determined by the modulation depth. At an m of 1 (fully modulated), the power of each sideband is half that of the f_c component. Figure 1.7 shows the waveforms and spectra of some example signals for comparison. At higher non-envelope AM rates ($>\sim$50 Hz), the sidebands are resolved on the basilar membrane and pitched tones are perceived (Viemeister 1979; Kohlrausch et al. 2000).

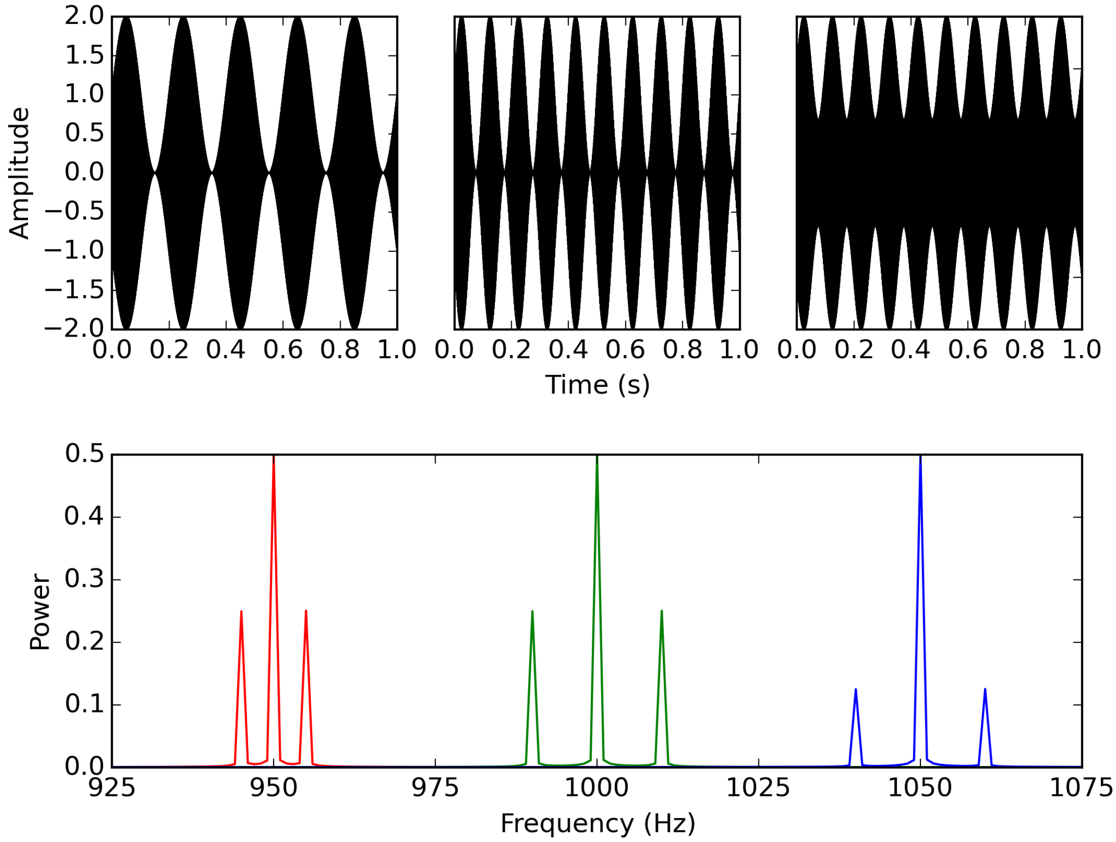

Figure 1.7: Upper: waveform representation of 3 sinusoidally amplitude modulated tone carriers. Lower: spectrum representation of 3 sinusoidally amplitude modulated tone carriers. Red: $f_c = 950$ Hz, $f_m = 5$ Hz, $m = 1$. Green: $f_c = 1000$ Hz, $f_m = 10$ Hz, $m = 1$. Blue: $f_c = 1050$ Hz, $f_m = 10$ Hz, $m = 0.5$.

1.7.2 Subcortical Central Auditory System

The central auditory pathway begins at the auditory nerve at ends at the auditory cortex. In the central auditory pathway, the electrical pulses generated at the auditory nerve are transformed by successive stages of specialised nuclei (Frisina 2001; Behler and Uppenkamp 2016). The anatomical basis of the subcortical central auditory pathway is well characterised. The afferent signals from the initial mechanotransduction by the organ of corti pass through the ascending structures of the cochlear nuclei, the superior olivary nuclei, the inferior colliculi (IC) and the medial geniculate bodies (MGB) of the thalamus (Grothe et al. 2010). Figure 1.8 shows a representation of these early auditory pathways. Due to the electrophysiological nature of the central nervous system, many parallels can be drawn between auditory processing and signal processing in the digital domain. Technological advances in computing and digital signal processing have allowed researchers to develop quantitative models of auditory function that can then be tested with physiological and psychophysical data (Lorenzi et al. 1995; Verhulst et al. 2018). A qualitative change in the way AM is encoded has been directly recorded in the central auditory pathway of non-human primates and other animals (Frisina 2001; Gao et al. 2016; Wang 2018). Further, it is proposed that separate subpopulations of neurons are responsible for each coding type. At lower levels of the pathway up to the auditory thalamus, the predominant coding type is an explicit temporal-based coding. This neuronal subpopulation show synchronised firing activity phase-locked to the temporal modulation of the auditory signal. At higher levels of the auditory pathway, activity from an unsynchronised neuronal subpopulation is predominant, and average firing rate is directly linked to the temporal structure of the input (Wang et al. 2003). The mechanisms for this transformation from synchronised temporal coding to unsynchronised firing rate coding are largely unknown but this is an active topic of research (Gao et al. 2016; Carbajal and Malmierca 2018).

Current models of human AM processing include a gammatone filterbank that models the cochlea and initial mechanotransduction (Patterson et al. 1987; Lyon et al. 2010). The output from this stage is then nonlinearly processed, typically modelled by half-wave rectification (or Hilbert transform) (Hudspeth 2008; Ainsworth et al. 2012). The signal is then passed to a modulation filterbank and a detector mechanism (Dau et al. 1997a; Joris et al. 2004; Xiang et al. 2013).

To summarise, the physiological architecture of much of the auditory pathway has been well-characterised. Robust anatomically-informed computational models of subcortical auditory processing have been developed. A progressive change in AM encoding type from explicit temporal coding to implicit rate coding can be measured in the central auditory pathway. Current theories of AM processing hold that a modulation filterbank best accounts for the observed physiological and psychophysical data, however, it is currently not known how or where such

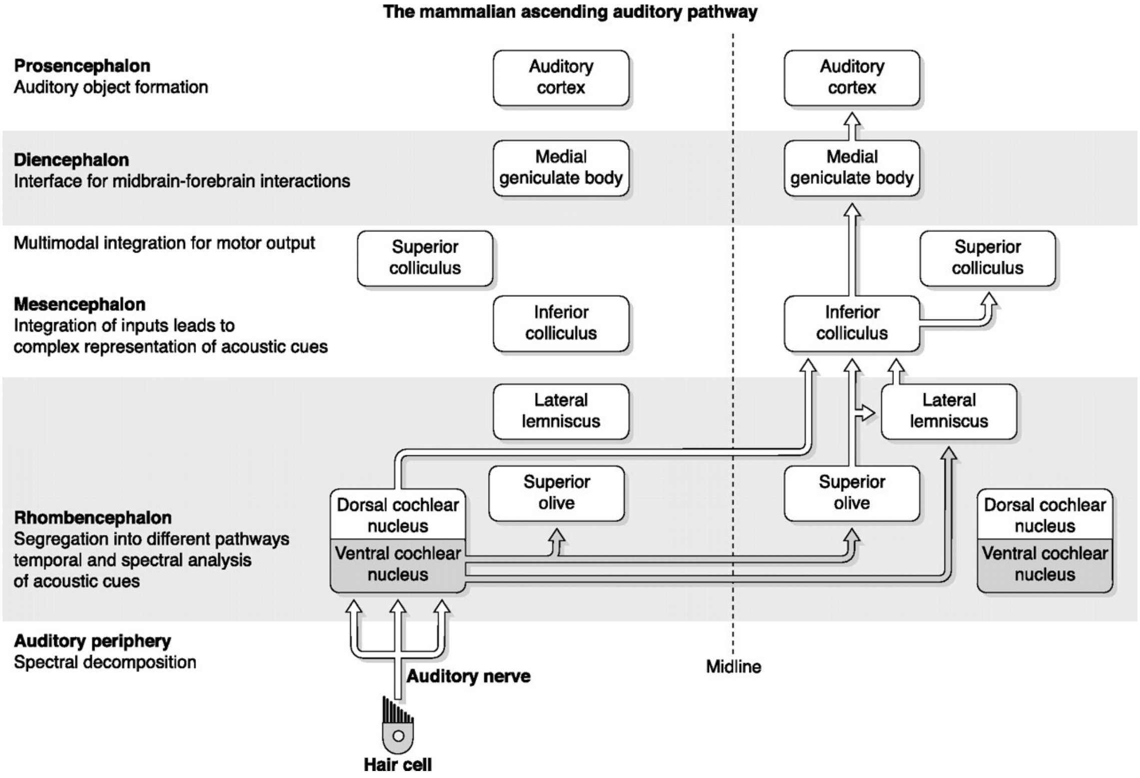

Figure 1.8: Simplified diagram of the ascending auditory pathway. Reproduced with permission from Grothe et al. (2010). Copyright 2010, American Physiological Society.

a system may be implemented.

1.8 Cortical Processing of Amplitude Modulation

Auditory signals are passed from the thalamus to auditory cortical regions such as Heschl's gyrus (HG), the location of primary auditory cortex (Abdul-Kareem and Sluming 2008) and peri-Sylvian regions such as the superior temporal gyrus (STG) (Grothe et al. 2010). Human auditory cortex is located on the supratemporal plane, within the Sylvian fissure, and comprises two thirds of the STG. It can be divided into three parts; the planum polare (PP), HG and the planum temporale (PT) (Moerel et al. 2014); see Figure 1.9. However, from two to six distinct cytoarchitectonic areas have also been described in secondary auditory cortex, reflecting the uncertainty surrounding the exact boundaries of these auditory areas (Cammoun et al. 2015). Haemodynamic activation to AM stimuli has been found within auditory cortex, specifically lateral HG, STG and PT (Giraud et al. 2000); see Figure 1.10 for fMRI clusters resulting from a contrast of 4 and 40 Hz AM noise over unmodulated noise. Further, fMRI activation in these regions has been shown to increase more to AM variation that to FM variation, suggesting a functional specialism of the neuronal ensembles (Hart et al. 2003).

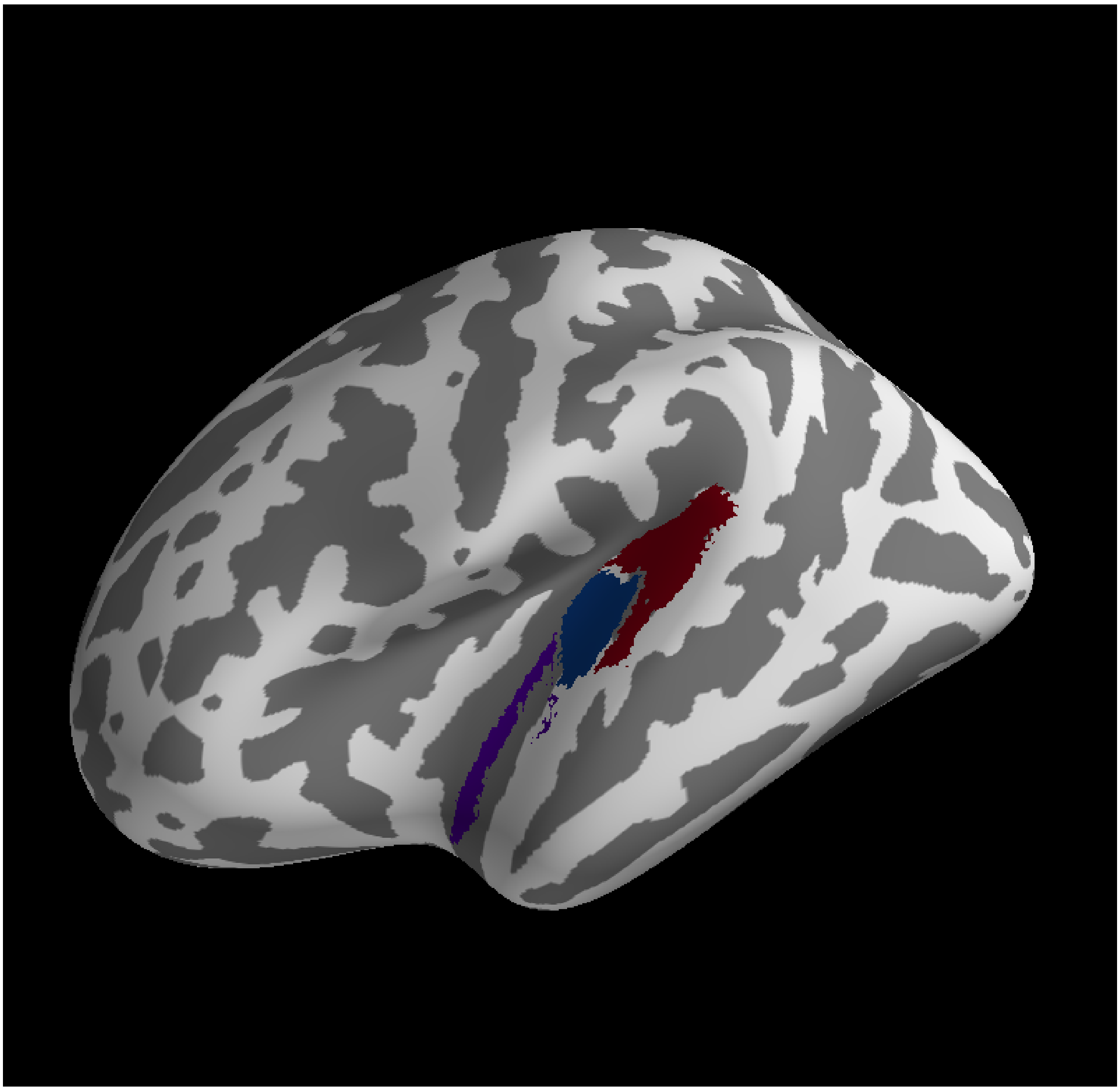

Figure 1.9: Location of human auditory cortex. Red: PT. Blue: HG. Purple: PP.

A change in encoding type as AM rate increases was first found using invasive single-neuron investigations in non-human primates (Lu et al. 2001; Wang

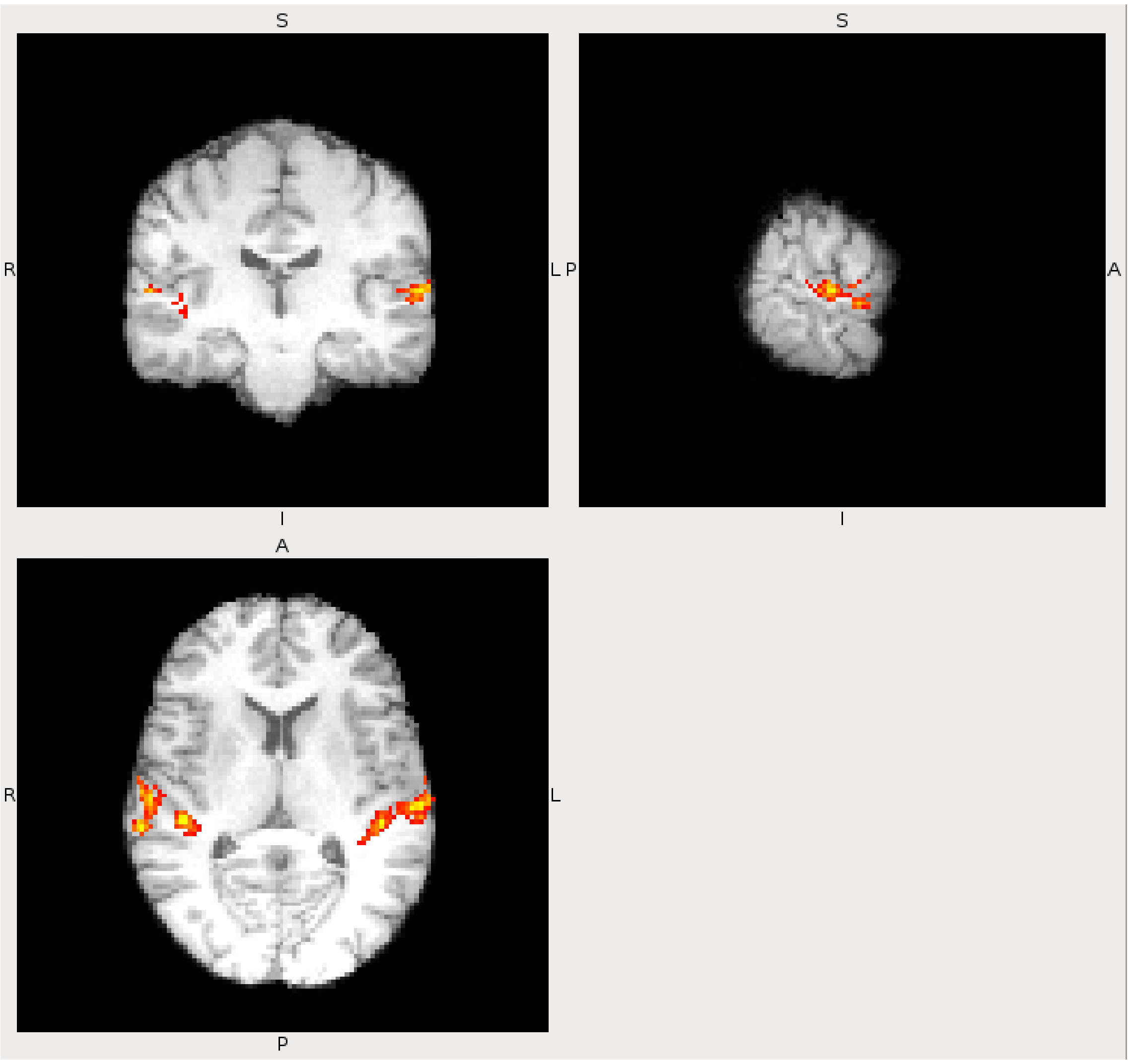

Figure 1.10: Clusters of activation in bilateral pSTG/PT in response to 4 and 40 Hz sinusoidally amplitude modulated broadband noise conditions (combined) over unmodulated noise.

et al. 2003) but evidence has also been found for this change in human cortex (Harms and Melcher 2002; Tang et al. 2016). Non-invasive human neuroimaging techniques measure the summed activity of a great number of neural populations, or indirectly index this through the BOLD signal, so it is notable that this organisation remains represented at this level. Harms and Melcher (2002) measured the BOLD response waveshape and amplitude while trains of noise bursts were presented with repetition rates from 1 – 35 Hz. In HG, response amplitude increased sightly with increasing rate, but the main effect was a change in waveshape from sustained to phasic; two distinct peaks in activity at train onset and offset. This encoding at the cortical level may reflect the perceptual change from many distinct noise bursts to one continuous noise period. This effect was absent in the inferior colliculus and started to emerge at the MGB, showing a systematic progression over these contiguous stages of the auditory system. In the IC, the response amplitude increased with increasing rate, but the waveshape remained the same; a sustained response for the length of the train (Harms and Melcher 2002). In the STG the observed effect changed in a subtle way, with responses to all rates showing an increasingly phasic pattern, though the distinction between high and low rates was still clear (Harms and Melcher 2002). Further work examined the effect of manipulating other sound features and found that the temporal envelope was most influential over fMRI activation waveshape and that sound level and bandwidth had no effect (Harms et al. 2005). This finding further emphasises the importance of the temporal envelope of the stimulus when processing AM sound.

It is well established that tonotopic organisation exists in the auditory system that persists from the cochlea throughout the auditory pathway up to the cortex. This direct representation of frequency information results in adjacent neuronal populations encoding adjacent frequencies across auditory cortex (Humphries et al. 2010; Saenz and Langers 2014; Moerel et al. 2014; Dick et al. 2017).

Figure 1.11 illustrates this tonotopic organisation. Direct topographic representation of stimulus features appears to be an underlying principle of sensory neural organisation and evidence has been found for periodicity (AM rate) on the orthogonal dimension to tonotopy in auditory cortex (Barton et al. 2012). The fMRI response was measured as broadband (0 – 8000 Hz) SAM noise was presented at different AM rates (from 2 – 256 Hz). Clear gradients were found and 11 auditory field maps were delineated, replicating the findings from animal studies. Similar spatial representations for AM rates that are approximately orthogonal to tonotopic maps have also been reported by Herdener et al. (2013). This has been supported by recent evidence from magnetoencephalography (MEG); a more direct measure of neural activity (Weisz and Lithari 2017). Group level results for tonotopic and modulotopic organisation can be seen in Figure 1.12. Further

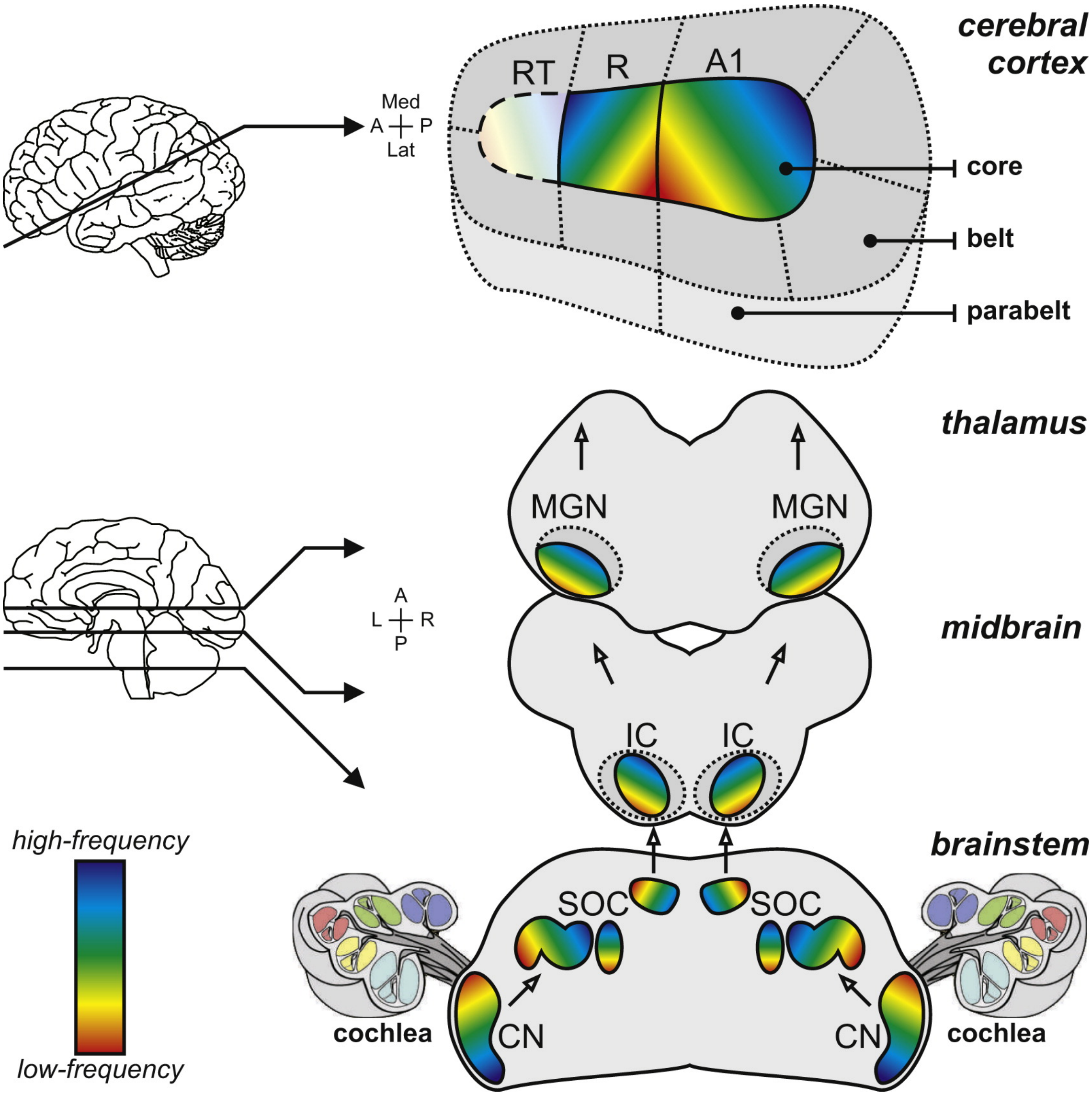

Figure 1.11: Tonotopic organisation throughout the auditory system. Reproduced with permission from Saenz and Langers (2014). Copyright 2013, Elsevier.

evidence for modulotopic organisation comes from an invasive electrophysiological study that found that sensitivity to AM rate changed along the length of the STG. They found that the anterior section activated more to high frequency fluctuations (with a relatively constant spectrum) whereas the posterior section activated more to slower fluctuations (with more spectral variation) (Hullett et al. 2016). Research investigating the representation of concurrent spectro-temporal modulations has suggested that these tono- and modulotopic maps may represent a tradeoff between spectral and temporal processing (Santoro et al. 2014).

To summarise, AM is clearly represented in auditory cortex, primarily mani-

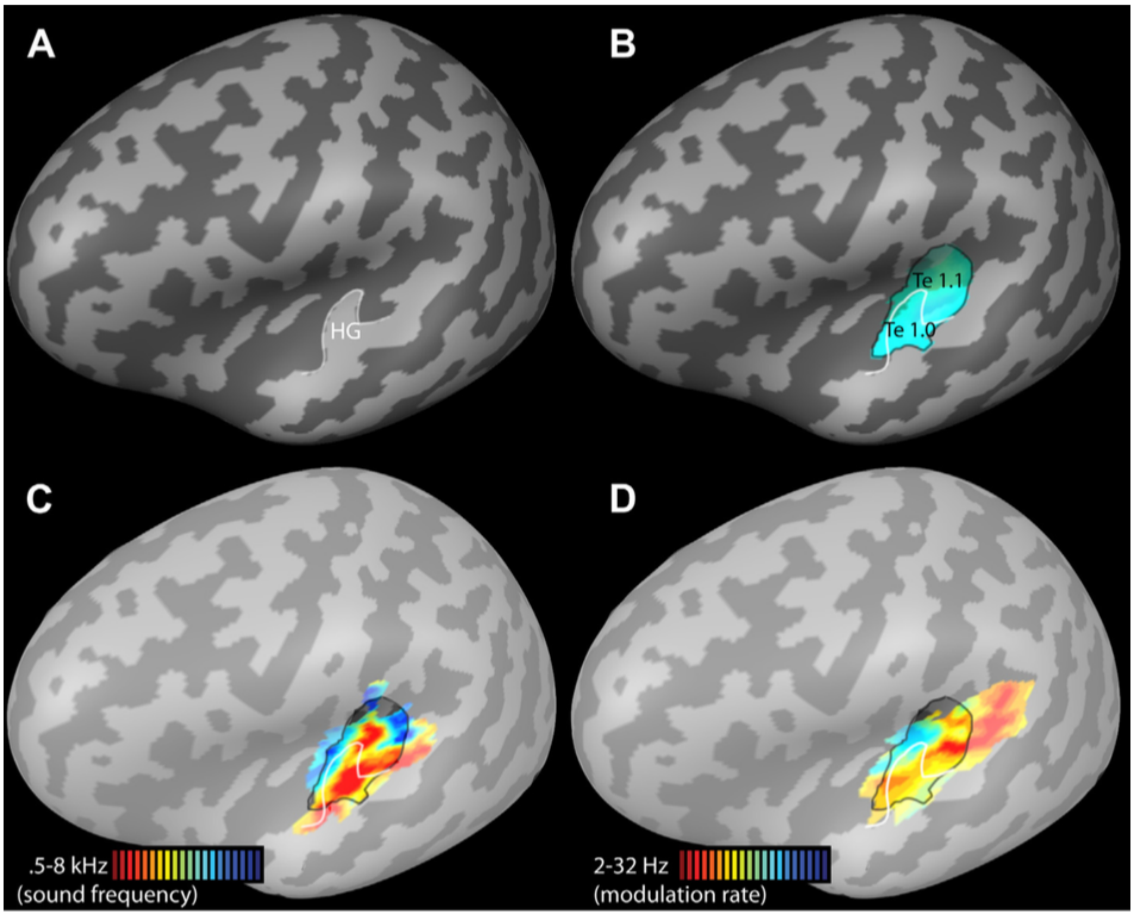

Figure 1.12: Tonotopic (C) and modulotopic (D) organisation in the cortex. Reproduced with permission from Herdener et al. (2013). Copyright 2013, Elsevier..

festing in the secondary auditory cortical regions STG and PT. The coding of AM rate appears to show a qualitative change at progressive stages of the central auditory pathway, and as AM rate increases, consistent with subcortical investigations. Tonotopic organisation is a key principle of the auditory system and some evidence has been found for an orthogonal modulotopic representation.

1.9 Cortical Asymmetries in Auditory Processing

It has been known for decades that a general specialism of the left hemisphere for speech processing and production exists (Berker et al. 1986). Since these initial findings, much progress has been made in modelling the systems underlying speech and language (Hickok and Poeppel 2007; Hickok 2012). Technological developments in structural and functional brain imaging have enabled detailed studies of the brain and gross structural hemispheric asymmetries have been found to exist in human cortex (Kong et al. 2018; Chiarello et al. 2016). Auditory and language areas have been specifically examined and these too show structural asymmetries (Abdul-Kareem and Sluming 2008; Moerel et al. 2014; Meyer et al. 2014; Galuske et al. 2000). In fact, auditory area PT has been described as the most asymmetrical cerebral structure in the whole brain (Prete et al. 2016). Investigations measuring structural connectivity have found that the arcuate fasciculus (the white matter bundle connecting Wernicke's and Broca's areas; key nodes in the language network) is more developed in the left than the right hemisphere, and this was related to behavioural differences (Catani et al. 2005, 2007). Another asymmetry of structural connectivity has recently been found where right auditory cortex was generally more integrated into the wider connectome than the left, particularly with respect to interhemispheric connectivity (Mišić et al. 2018). This result is consistent with functional network connectivity investigations and provides evidence that this structural asymmetry is associated with function (Andoh et al. 2015). Functional asymmetries of auditory areas have also been found, such as an asymmetry of sensitivity to temporal structure, with slow temporal modulations preferentially driving the right hemisphere (Boemio et al. 2005). Conversely, evidence has been found for preferential activation of the left hemisphere by rapid temporal modulations, such as those found in speech (Zaehle et al. 2004; Abrams et al. 2008). Multiple behavioural and imaging studies have provided evidence for functional asymmetries (Millman et al. 2011; Saoud et al. 2012; Liem et al. 2014; Han and Dimitrijevic 2015; Hugdahl 2011; Hugdahl and Westerhausen 2016; Schremm et al. 2018). Several theories have been developed to explain these asymmetries and concurrently account for the empirical findings. One prominent theory holds that the left hemisphere is specialised for processing temporal aspects and the right hemisphere is specialised for processing spectral aspects of sound (Zatorre and Belin 2001). In this study metabolic activity in the brain (measured with positron emission tomography) was compared when stimuli that differed in either spectral or temporal content were presented. The responses showed that activation to rapid temporal changes was weighted towards the left hemisphere while activation to spectral changes was weighted towards the right hemisphere. This theory was later extended to explain music processing as primarily governed by the right hemisphere (Zatorre et al. 2002).

Alternatively, the Asymmetric Sampling in Time (AST) hypothesis proposes that structural asymmetry of auditory cortical regions is due to functional specialisms. It proposes that auditory cortex in each hemisphere preferentially processes aspects of a speech signal based on sampling windows of different lengths. As the brain has limited computational power, it must sample the continuous analogue sound signals received by the ear as part of the encoding process. The AST hypothesis accounts for the asymmetry by proposing that the right hemisphere samples and integrates information over 200–250 ms time windows, corresponding to ∼4 Hz and the left hemisphere samples and integrates information on much shorter timescales, ∼25 ms, corresponding to ∼40 Hz. Figure 1.13 shows how sampling windows at these lengths roughly correspond to a speech signal (Poeppel 2003) This is a simplistic representation, however, as there is evidence for more optimised sampling where the timing of these sampling windows is adjusted based on acoustic properties of the signal (Doelling et al. 2014). The AST hypothesis was informed by observed oscillatory patterns in the theta (4–8 Hz) and low gamma (∼40 Hz) ranges during speech perception and there is mounting evidence for the role of this activity (Poeppel 2003; Giraud et al. 2007; Peelle and Davis 2012; Giraud and Poeppel 2012; Obleser et al. 2012; Luo and Poeppel 2012). This temporal sampling account may reflect a more general principle underlying speech processing as suboptimal temporal sampling is thought to underpin language deficits in some disorders, such as dyslexia (Goswami and Leong 2013; Cutini et al. 2016) and autism spectrum disorder (O'Connor 2012). It is widely accepted that speech is processed asymmetrically and that extraction and tracking of the speech amplitude envelope is a key component of this mechanism (Kubanek et al. 2013; Ghitza et al. 2013).

Despite the allure of these left-right dichotomies, it is likely that these ac-

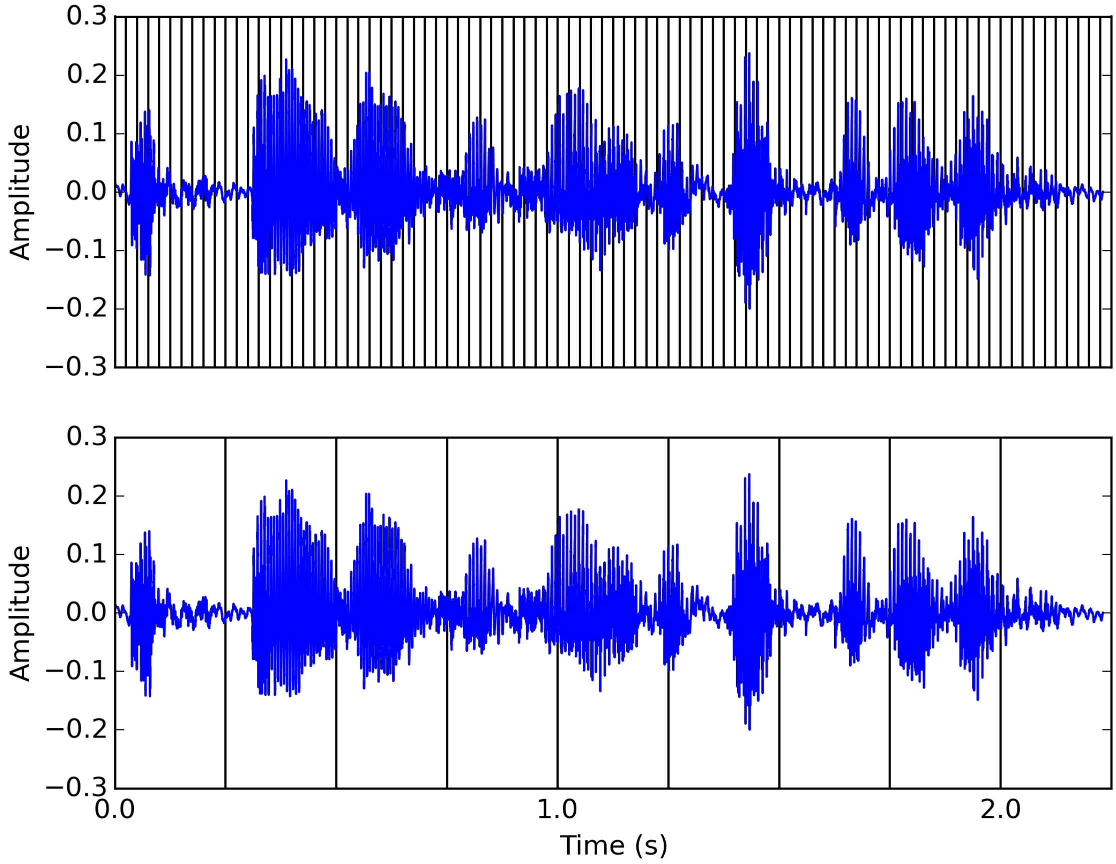

Figure 1.13: Upper: representative sampling windows of 25 ms (40 Hz). Lower: representative sampling windows of 250 ms (4 Hz). The speech sample is the same as in Figure 1.1.

counts are too reductionist; reflecting the general trend in neuroscientific theory (Krakauer et al. 2017). Though their parsimony is compelling, it has been argued that they may not fully account for known properties of the speech signal (McGettigan and Scott 2012). An alternative account suggests that, on closer inspection, there may be a specialism of the right hemisphere for sounds that are longer or change more slowly but there may be no temporal specialism of the left hemisphere (Scott and McGettigan 2013). Further, there is contrasting evidence for speech segmentation using ∼40 Hz sampling windows, as shown by time-reversion studies and the presence of phonetic components longer than the proposed window (Saberi and Perrott 1999; Ueda et al. 2017; McGettigan and Scott 2012).

To summarise, there is substantial evidence for hemispheric asymmetries of the auditory cortical regions. These asymmetries are both structural and functional, and interhemispheric connectivity may be a crucial underlying component. Multiple accounts have been proposed to explain these asymmetries but there is no current consensus on the exact role they play in speech processing.

1.10 Thesis Aims and Plan

Behavioural and neuroimaging research has revealed much about how AM and speech are processed by the auditory system, however, there are still unanswered questions. Advances in understanding are often made through the use of novel experimental techniques. There is a paucity of research combining non-invasive neuroimaging and neurostimulation methods to investigate these questions. Specifically, understanding AM processing is a promising, novel application for fMRI-guided TMS. Currently, it is not known if it possible to modulate human AM processing ability using fMRI-guided TMS. Further, a wide range of TMS protocols has been developed and widely different responses have been elicited. This proliferation of pulse protocols has resulted in TMS becoming a very versatile tool, but much further research is required to fully understand the effects of each protocol and the extent to which they transfer across sensory modalities. It is not currently known how effective different TMS protocols are at modulating AM processing. Finally, fMRI-guided TMS has the potential to inform us about speech processing more broadly. Some contrasting perspectives on the functional asymmetry of speech processing in the brain have been presented. The combination of fMRI and TMS allows precise targeting of neural disruption, which can be systematically applied to each cortical hemisphere in order to investigate these proposed asymmetries. However, the effectiveness of fMRI-guided TMS for this higher order task has also not previously been explored.

This investigation was designed to address a set of overarching research questions.

- Firstly, is fMRI-guided TMS an effective method for modulating AM processing?

- Secondly, does the effectiveness of TMS for modulating AM processing differ based on the TMS protocol used?
- Finally, is fMRI-guided TMS an effective method for further understanding the functional asymmetry of speech processing?

Three combined psychophysical, neuroimaging and neurostimulation experiments were designed and conducted to address these research questions. This introduction is followed by a methods chapter which briefly introduces the diverse range of experimental methods. This takes the form of a general introduction to the methods and then addresses specific challenges that auditory neuroimaging and neurostimulation research presents. The core of the thesis is three experimental chapters in the form of research articles. Each experimental chapter thoroughly describes a novel empirical study conducted to address aspects underlying these research questions. The first experimental chapter (Chapter 3) describes a study that combined a behavioural measure of AM detection with fMRI-guided online TMS. The second experimental chapter (Chapter 4) describes a study that combined a behavioural measure of AM detection with fMRI-guided TMS applied while the underlying state of the cortex was manipulated using behavioural selective adaptation to AM noise. The third experimental chapter (Chapter 5) describes a study that combined a behavioural measure of AM depth discrimination with an fMRI-guided offline TMS protocol; continuous theta-burst stimulation. The experimental chapters are followed by a general discussion of the core findings, how these relate to previous research, and how they can be used to inform future investigations.

Chapter 2

Methods

This chapter provides a brief outline of the three primary experimental methods used in this thesis. The three methods; psychophysics, functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) have all been previously applied to investigate auditory processes, however this often raises methodological challenges. Methodological developments have ameliorated the effects of these challenges, however, there are still some limitations on the investigation of auditory function. Fully understanding the strengths and limitations of each method allows researchers to apply them in complementary ways and gain truly unique insights into how the brain processes sound and speech. More detailed technical information regarding the techniques as applied in this thesis are provided in the relevant experimental chapters.

2.1 Psychophysics

Psychophysical techniques measure the relationship between the level of a stimulus feature and perceptual experience of the stimulus. They are behavioural techniques that are commonly used to make inferences about the internal processes underlying perception. It is possible to separate and assess the contribution of different processes by systematically modifying stimuli in a single dimension and measuring the effect on perception. In its simplest experimental form, participants are told to respond when they detect a stimulus (e.g. hear a tone) and the researcher systematically adjusts the intensity to find the point where the participant can just detect it. Although more complex automated procedures are now common, this general technique has remained remarkably constant since it was established in the 19th century (Read 2015). This simple, powerful technique has made substantial contributions to our understanding of sensory perception and the neural circuits underpinning behaviour and continues to do so.

Parsimonious models of the data-generating process underlying the relationship between stimulus intensity and perception have been developed. A monotonic, non-linear relationship is often used to model many sensory processes and this can be represented by the psychometric function (Wichmann and Hill 2001). The psychometric function describes the probability of a correct response as a function of stimulus intensity and can be used to estimate summary statistics such as the threshold or slope. Equation 2.1 describes the basic form of the 2-parameter psychometric function.

$$P_a(x) = \Psi_{\alpha(a),\beta(a)}(x) \tag{2.1}$$

Ψ describes the kernel type of the function. The exact form of the psychometric function can be altered by changing this parameter and Gaussian or Weibull cumulative distribution functions are often used (Wichmann and Hill 2001; May and Solomon 2013). Psychometric functions with these kernels have a sigmoid shape with a constant slope about the point of subjective equality (PSE: 50% threshold for 2-alternative forced choice tasks). This sigmoidal relationship is similar across many sensory modalities, suggesting a general principle underlying sensory representation. α and β are free parameters that represent the threshold and slope values when a Gaussian kernel is used. The threshold is the response probability that corresponds to a given stimulus value. For example, the 50% threshold is the value of x that corresponds to the PSE (for a 2-AFC task), however, any threshold value can be estimated. Estimating thresholds is useful in behavioural studies where you expect an experimental manipulation to affect detection of a stimulus. The threshold can be determined and then the task can be performed at threshold; a deviation from performance at the chosen threshold percentage can indicate an experimental effect. Alternatively, psychophysical data can be collected and a threshold estimated, then an experimental manipulation can occur, followed by another psychophysical task and threshold estimate. This method requires a direct comparison of estimated thresholds to determine the presence of an experimental effect. The slope parameter determines the steepness of the linear part of the psychometric function and represents the change in y units for each change in x. More advanced models can also account for the guess rate and lapse rate, represented by γ and λ parameters respectively. γ describes the likelihood of a correct response if no stimulus is presented; this baseline chance level varies based on the choice of psychophysical task and can be calculated as the reciprocal of the number of alternatives; i.e. $1/n$ for an n-AFC task. λ describes the expected proportion of trials where the participant has an attentional lapse and makes a guess response (Wichmann and Hill 2001; Prins 2012). These two parameters constrain the psychometric function in the x direction so that the lower asymptote is equal to γ and the upper asymptote is equal to $1 - \lambda$. Equation 2.2 describes the 4-parameter psychometric function.

$$\Psi(x; \alpha, \beta, \gamma, \lambda) = \gamma + (1 - \gamma - \lambda) F(x; \alpha, \beta) \tag{2.2}$$

An example of a 3-parameter psychometric function with a γ parameter of 0.5 and cumulative Gaussian kernel can be seen in Figure 2.1.

Although psychometric functions are commonly estimated using a parametric model and (constrained) maximum likelihood, alternative approaches include

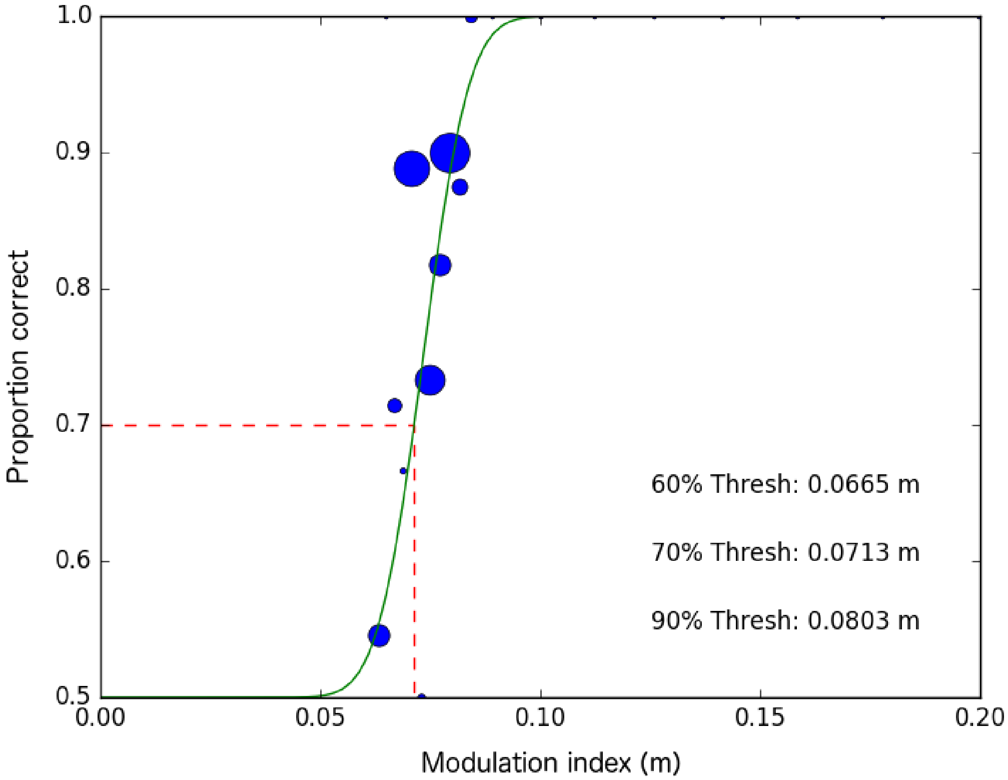

Figure 2.1: An example 3-parameter psychometric function (green line) for a 2AFC AM detection task. The size of the blue points indicate the number of trials at that modulation depth. Estimated thresholds are given for 60%, 70% and 90% thresholds. The red dashed line indicates 70% accuracy.

Bayesian optimisation (Schütt et al. 2016; Watson 2017) and model-free variants (Zchaluk and Foster 2009). An example of a Bayesian psychometric function and posterior distributions for three parameters, estimated with Hamiltonian Monte Carlo sampling can be seen in Figure 2.2.

2.1.1 Auditory Psychophysics

There is a long history of applying psychophysical techniques in the auditory domain. The most common auditory psychophysical procedure is the pure tone audiogram, which is routinely used to assess audiometric hearing thresholds in clinical practice (and in the experiments performed as part of this thesis). This procedure measures the smallest intensity required for a listener to hear a sinus-

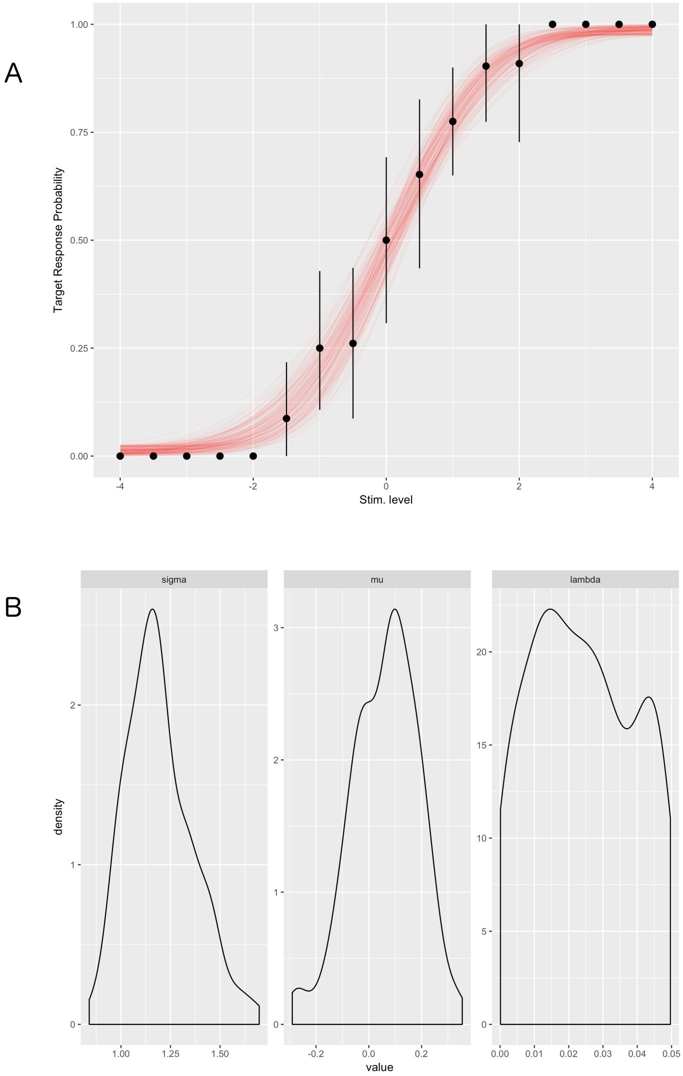

Figure 2.2: (A) Bayesian psychometric function. (B) Posterior distributions for three parameters representing α, β and λ.

oidal pure tone by decreasing the intensity until the tone is no longer perceived. The point at which the tone is just audible is considered the threshold of hearing for that frequency. This procedure can then be repeated at many audio frequencies to map the response for the full frequency range of the auditory system. An example of a pure tone audiogram can be seen in Figure 2.3. In the case of the pure tone audiogram, the parameter of interest is sound level (as a function of tone frequency), however this general procedure can be modified and applied to investigate many aspects of auditory processing.

For example, by altering the AM depth of controlled noise stimuli and meas-

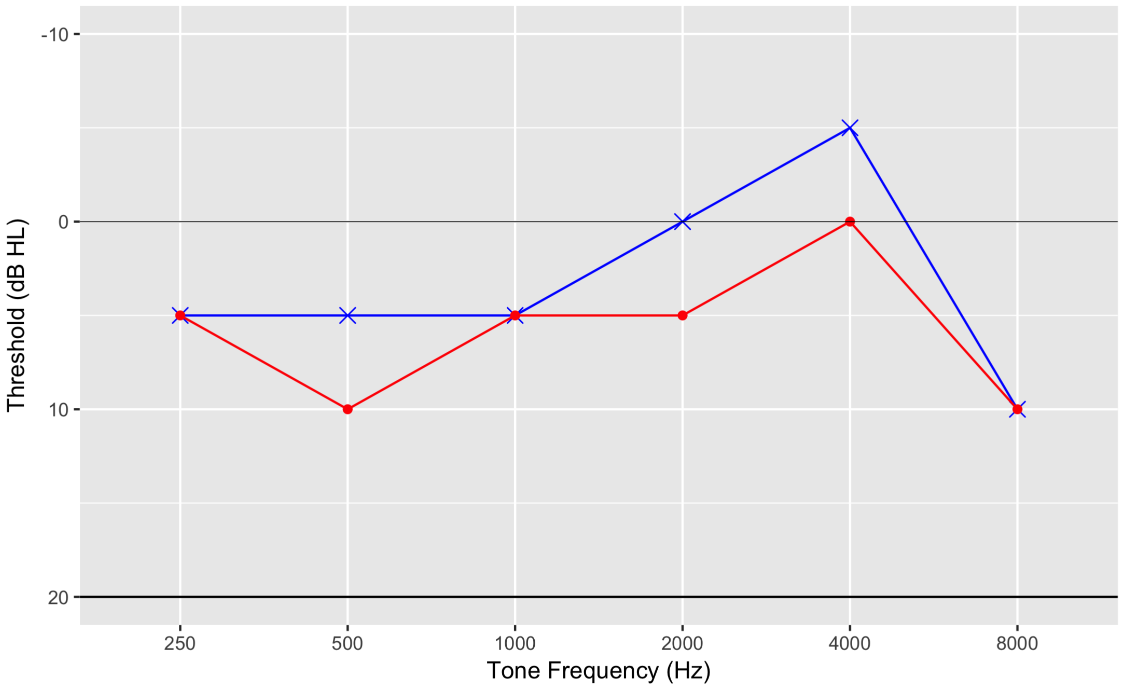

Figure 2.3: An example pure tone audiogram. The blue points are measured from the left ear and the red points are measured from the right ear. The thin black line indicates normal hearing thresholds and the thick black line indicates the cutoff for minor hearing loss.

uring how detection changes, it is possible to probe the sensitivity of the auditory system to this specific aspect. This is how the temporal modulation transfer functions (TMTFs) described in the introduction chapter were measured. To recap, the TMTF is a measure of the relationship between AM rate and AM detection threshold (Rodenburg 1977; Viemeister 1979). In these studies, 2-AFC behavioural tasks and an adaptive staircase tracking procedure were used to identify the 70.7% detection threshold for AM at rates from 2 – 4000 Hz. This task contained two periods of 500ms broadband noise, one of which was amplitude modulated. The participant had to identify which period was modulated and the depth of modulation was varied from trial to trial based on an adaptive (1-up, 2-down) rule (Levitt 1971). This simple adaptive procedure results in the task becoming more difficult (AM depth decreases), when participants respond correctly on two consecutive trials and the task becoming easier (AM depth increases) when participants respond incorrectly on a single trial. The specific sequence of step sizes was chosen to hone in on the 70.7% threshold value. AM depth discrimination tasks such as those conducted by Wakefield and Viemeister (1990), have a similar structure. In these 2AFC tasks, both intervals have been amplitude modulated, however one period has a greater modulation depth. A central period was added that was always unmodulated to provide a reference for the listener. The participant had to identify which interval contained amplitude modulation at a greater depth. Figures showing examples of AM detection and AM depth discrimination TMTFs can be found in the previous chapter.

2.2 Non-invasive Neuroimaging & Neurostimulation

Non-invasive neuroimaging and neurostimulation methods are invaluable tools for investigating how brain structures and networks are linked to behaviour. These overarching terms describe a number of relatively recently developed methods that are remarkably distinct in the variety of biological signals they measure (or initiate through stimulation) and how these signals are acquired. This leads to techniques with varying properties, such as temporal and spatial resolution, that can often be combined in complementary ways. Although many of these techniques have desirable properties for investigating the AM processing of sounds, this section is focussed on the methods used in this thesis.

2.2.1 Functional Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is commonly used to produce high resolution visualisations of the brain. These structural images are used in clinical contexts, for example, to assess the presence of lesions or in an experimental context to investigate how the brain is structured. This method exploits the phenomenon of magnetic resonance combined with targeted radio frequency (RF) energy to non-invasively acquire images with high spatial resolution without the use of ionising radiation (Plewes and Kucharczyk 2012). When protons (hydrogen atoms in water or fat within the brain) are exposed to a strong magnetic field (B_0), their spins align with it; there is however, a slight mismatch between the number of protons aligned in phase and anti-phase (spin, in this context, refers to the physical property analogous to spinning on an axis). This results in a small net alignment with B_0 and this mismatch is proportional to the field strength of B_0 (leading to higher signal to noise ratios in the resulting images). Magnetic gradients are applied to modify B_0 strength across the two planes orthogonal to B_0, which alters the resonance frequencies of protons based on their location by the gyromagnetic ratio. By applying RF energy at specific resonance frequencies, it is possible to 'tip' the axis of selected slices of protons away from B_0. When RF energy transmission ceases, these protons then 'relax' back into alignment with B_0 and the energy that this process releases is measured. It is possible to acquire images with different contrasts by exploiting the different timing properties of protons in different molecules, for example longitudinal relaxation time (T_1-weighted) or transverse relaxation time (T_2-weighted).

This is the basis of how structural MR images are acquired, however, magnetic resonance is a versatile phenomenon that can also be exploited to investigate the functional relevance of different brain regions and networks. This *functional* magnetic resonance imaging (fMRI) allows researchers to measure which brain regions or networks are active during specific behaviours. Blood-oxygenation level-dependent (BOLD) fMRI exploits the different magnetic properties of oxygenated and deoxygenated blood to detect and localise changes in blood flow to brain regions (Kwong et al. 1992; Logothetis et al. 2001). When a brain region is active, the local neurons require more oxygen and a surge in blood flow to that area is experienced. By measuring and contrasting the BOLD signal in brain regions or networks during behaviour, researchers can infer what brain regions or systems are active during specific behaviours. This response is an indirect measure of neuronal activity (Logothetis et al. 2001). This haemodynamic response can be sampled using fMRI and the data analysed to locate the voxels of origin. Through a combination of structural and functional imaging, researchers are able to localise and investigate auditory functions in novel ways. Despite the unique insights that MRI can provide, investigating auditory function with MRI presents with a set of methodological challenges (Peelle 2014). Most importantly, the scanner environment is very loud (∼127 dB SPL, for a typical sequence). This scanner noise can mask auditory stimuli and lead to significantly lower measures of activation (Elliott et al. 1999). It can also affect non-auditory processes and the psychological effects can confound measures of brain activity (Andoh et al. 2017). This problem led to the development of sparse imaging sequences that aim to minimise the impact of the scanner noise on data quality (Hall et al. 1999; Müller et al. 2003; Zaehle et al. 2004; Talavage and Hall 2012). One such sequence, interleaved silent steady-state (ISSS), exploits the sluggishness of the haemodynamic response so that the auditory stimuli can be presented during a 'silent' period followed by a regular acquisition period when the response is sampled (Schwarzbauer et al. 2006). A graphical representation of an ISSS sequence can be seen in Figure 2.4. ISSS imaging significantly alleviates the noise masking problem and allows auditory experiments to be conducted with minimal adverse effects on data quality. Another problem that auditory researchers face is the large amount of anatomical variability in auditory areas between participants and hemispheres (Abdul-Kareem and Sluming 2008). Due to this variability, regions of interest (ROI) defined on a standard brain map are rarely adequate for comparative analyses across participants and auditory researchers must typically define each ROI on a per-participant basis.

2.2.2 Transcranial Magnetic Stimulation (TMS)

TMS is a non-invasive neurostimulation technique used to systematically modulate ongoing neural activity. A time-varying magnetic field is induced through a coil placed on the scalp and re-induces electrical current in the underlying neural tissue (Barker et al. 1985). A simulation of the induced current from a TMS pulse can be seen in Figure 2.5. It was initially applied in human motor physiology to

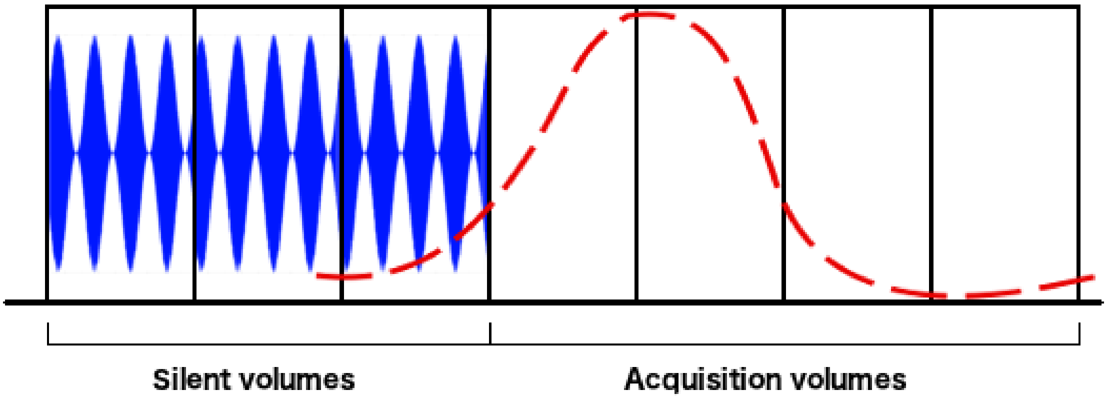

Figure 2.4: A representation of the ISSS sequence. Time is represented from left to right. Each box represents a volume and the blue envelope represents the auditory stimulus. In this example, three 'silent' non-acquisition volumes are followed by four acquisition volumes. The red dashed line represents the haemodynamic response.

probe cortico-spinal pathway functioning but has since been successfully applied in many psychological contexts (Klomjai et al. 2015). Causal inferences about the relationship between brain regions and behaviour can be made as neural activity is being directly manipulated. In its simplest form, single pulses of TMS can be delivered in order to temporarily disrupt cortical processing. This online protocol has been conceptualised as creating 'virtual lesions', akin to natural lesions found in neurological disorders such as stroke. However, the safe and temporary nature of these 'virtual lesions' allow investigations of brain function in neurotypical individuals. With this protocol, TMS is applied to areas of the cortex ostensibly involved in a behavioural task and performance is measured and compared to a baseline score. The cortical region is assumed to be critical for the task if performance is modulated by TMS (Pascual-Leone et al. 2000). Experimental manipulations such as altering the timepoint of stimulation allows researchers to probe the functional role of cortical regions in novel ways.

In addition to single pulse protocols, more complex TMS protocols have been developed and TMS has been shown to be a versatile technique. For example, a subthreshold conditioning pulse has been shown to affect the motor response to a suprathreshold pulse when the interstimulus interval is short (1 – 6 ms) (Kujirai et al. 1993). This short interval intracortical inhibition is just one of the many paired-pulse protocols developed to investigate human neurophysiology. In other studies, a train of multiple TMS pulses have been applied and the direction of effects have been shown to differ based on the specific frequency of pulse delivery (Klomjai et al. 2015). Low frequency (<1 Hz) stimulation typically shows inhibitory effects whereas high frequency stimulation (>5 Hz) typically shows excitatory effects. The effects of these repetitive TMS (rTMS) protocols can often outlast the period of stimulation, a property desirable for their use in clinical investigations. A commonly used complex rTMS protocols is continuous theta-burst stimulation (Huang et al. 2005). This protocol involves delivering bursts of 3 pulses at 50 Hz rate, each separated by continuously in a 20 or 40 second train (producing a total of 300 or 600 pulses). The efficiency of this protocol makes it preferable to some other rTMS protocols using lower frequencies of stimulation. A graphical representation of cTBS can be seen in Figure 2.6. This protocol is designed to imitate the effects of long-term depression at the neural level and has shown decreased M1 excitability after-effects lasting up to one hour post-TMS (Huang et al. 2005).

TMS can be applied to investigate auditory processes, however there are some unique methodological challenges. A loud click is emitted during TMS pulse delivery and this auditory artefact can lead to masking of auditory stimuli. As a consequence of this, offline rTMS protocols are particularly well-suited for investigating auditory processing. rTMS can be delivered and then auditory tasks can be completed in the subsequent (silent) aftereffect period. Thus, offline protocols like cTBS allow TMS effects to be measured when there is no concurrent artefact and auditory stimuli can be presented without interference.

Combining non-invasive neuroimaging and neurostimulation techniques is a useful way of gaining novel insights into many aspects of brain function. It is possible to apply the techniques in complementary ways by utilising the strengths

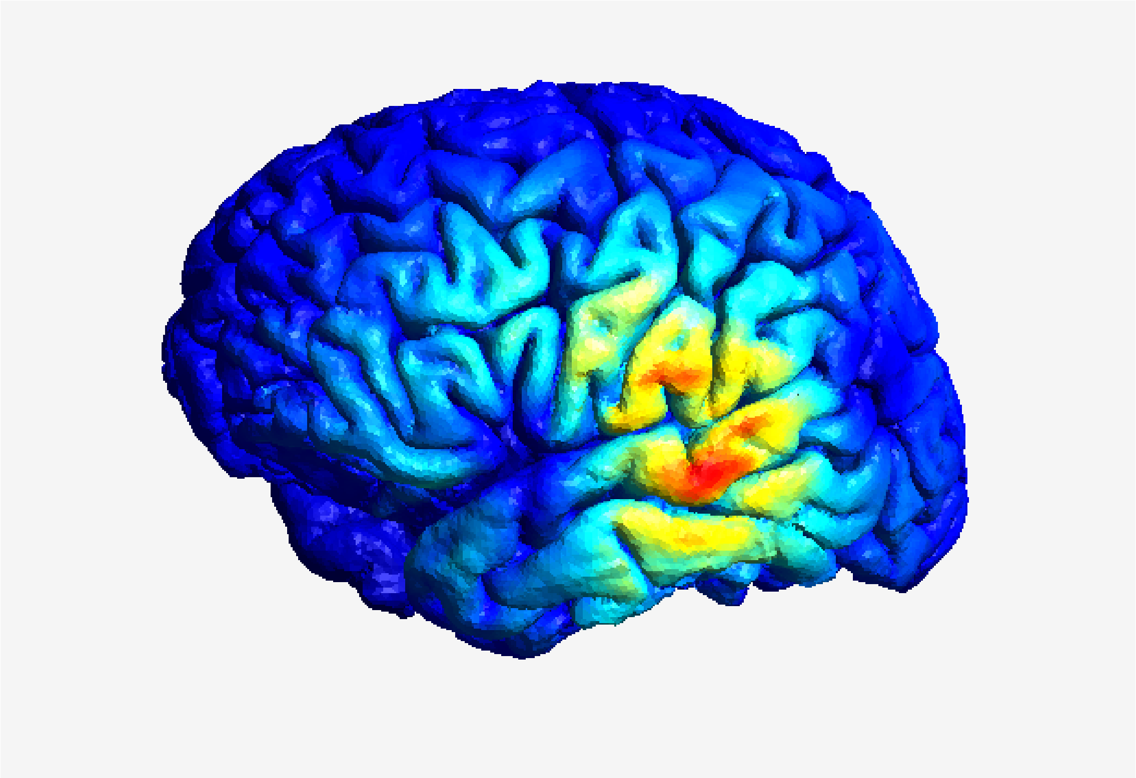

Figure 2.5: A visualisation of simulated current flow induced from a TMS pulse to the superior temporal gyrus using a 70mm figure of 8 coil.

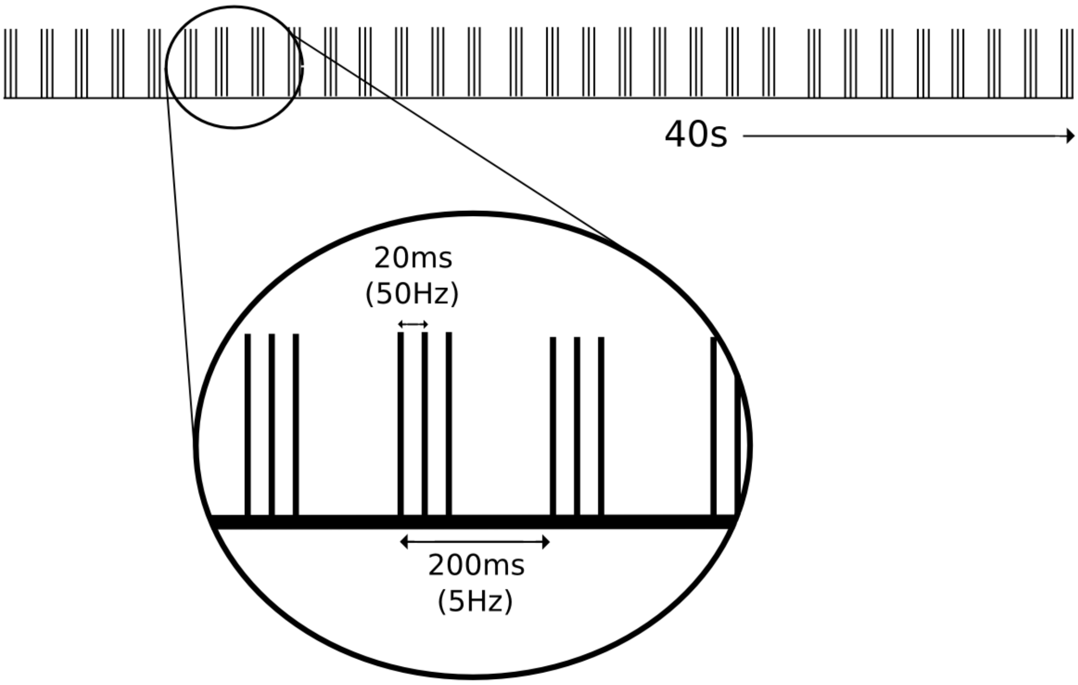

Figure 2.6: A schematic representation of the continuous theta-burst stimulation protocol. Each vertical line represents a single TMS pulse.

of each technique. This multimodal approach can improve existing methods or provide novel evidence about brain function from multiple sources. A common way of using MRI data to improve TMS praxis is through the use of MRI or fMRI neuronavigation (Sack et al. 2008). Neuronavigation utilises the high spatial resolution of MRI to improve the accuracy of TMS delivery. Structural MRI data can be used to ensure accurate targeting of cortical regions with TMS. This can be acquired for each study participant to minimise targeting variability due to anatomical differences and the exact location and trajectory of each pulse can be recorded to increase methodological rigour. Functional MRI data can also be incorporated to enable the accurate targeting of novel TMS sites. fMRI data acquired while an individual is performing the task of interest can be used to find peak locations of activation for that specific task. These locations can then be used as targets in a TMS study investigating this behaviour. Though technically challenging, it is also possible to apply TMS and MRI concurrently and these investigations provide the opportunity for unique insights into the haemodynamic effects of TMS during and directly following stimulation (Bestmann et al. 2005; Ruff et al. 2006, 2009).

To summarise, it is possible to investigate auditory processing using multiple behavioural, neuroimaging and stimulation methods. Auditory research poses some unique methodological challenges, primarily because many neuroimaging and stimulation methods are relatively loud. Methodological developments have allowed researchers to overcome the limitations of standard neuroimaging and stimulation techniques and apply these powerful methods to investigate auditory processes in novel ways. Many of these developments are relatively novel and there is still much scope for unique insights about the auditory system from their application.

Chapter 3

fMRI-guided Online Transcranial Magnetic Stimulation can Disrupt Detection of Low Frequency Amplitude Modulation

3.1 Abstract

Speech perception is a complex process and concurrent processing at multiple temporal scales is a key component. Amplitude modulations (AM) of the speech envelope are crucial for intelligibility but how they are processed is not fully understood. There is growing evidence for the functional role of hemispheric asymmetries of auditory cortical regions. To explain these functional asymmetries, auditory cortical areas in left and right hemispheres have been proposed to use temporal sampling windows of different lengths. Online transcranial magnetic stimulation (TMS) may be an effective and novel tool for investigating this. An at-threshold AM detection task with a broadband noise carrier was used to measure the influence of TMS and interleaved silent steady state functional magnetic resonance imaging was used to target the TMS. AM detection was measured for rates of 4 and 40 Hz with TMS to the left hemisphere, right hemisphere or a sham control. It was found that online TMS can be an effective tool for modulating sensitivity to AM. The detection of slow (4 Hz) amplitude modulations was impaired by TMS to auditory cortex in the right hemisphere only. This study provides causal support for the role of right posterior superior temporal gyrus/planum temporale in processing slow amplitude modulations. Interpretation of the 40 Hz results is more complex due to psychophysical thresholding concerns. These results are discussed and recommendations are made for future investigations.

3.2 Introduction

Amplitude modulations (AM) are an integral part of most natural sounds (Joris et al. 2004). The ability to detect amplitude modulation is vital for human speech communication as these fluctuations often provide linguistic information. The slowest amplitude modulations of speech, ranging from about 2 – 50 Hz, are categorised as the speech envelope (Rosen 1992) and low frequency (<16 Hz) AMs are particularly crucial for speech intelligibility (Drullman et al. 1994a,b; Greenberg 2004). Further, speech that has been severely degraded in the frequency domain can be understood with high accuracy provided that envelope information within a number of frequency bands is adequately preserved (Shannon et al. 1995). The human auditory system is also particularly sensitive to AM in this range; with peak sensitivity at ∼4 Hz, further suggesting behavioural significance (Viemeister 1979; Rees et al. 1986; Kohlrausch et al. 2000). The sensitivity of the auditory system to temporal modulations has been characterised as low-pass in shape, with sensitivity decreasing as rate increases above ∼4 Hz (Viemeister 1979). However, a reexamination of the literature has suggested that it may in fact be band-pass in shape, with a pass band of ∼2 – 5 Hz (Edwards and Chang 2013).

The full mechanism by which the central auditory system extracts amplitude modulation from an auditory signal is currently unknown. A key step in understanding this mechanism is elucidating the brain regions involved. Neuroimaging methods such as functional magnetic resonance imaging (fMRI) have proved particularly useful for this task (e.g Giraud et al. 2000). fMRI studies have shown that activation to AM is found throughout the auditory pathway, from the inferior colliculus, through the MGB of the thalamus, up to the cortex, however the response type qualitatively changes (Harms and Melcher 2002). Ascending stages of the auditory pathway show preferential responses to lower rates of AM and the shape of the response changes from phasic to sustained as the AM rate decreases (Harms et al. 2005). This account is supported by both invasive and non-invasive electrophysiological evidence (Nourski 2017; Teng et al. 2017). Such a graded response profile is consistent with a mechanism where the incoming signal is filtered and amplitude information is extracted by specialised brain areas at progressively lower rates, such as a cascading bank of modulation filters (Dau et al. 1997a; Giraud et al. 2000; Jepsen et al. 2008). It is commonly accepted that AM is extracted by a bank of filters, however it is not clear where this may be instantiated in the auditory system, although neurophysiological evidence has suggested it may be at the cortical stage (Xiang et al. 2013).

Whilst neuroimaging methods have allowed researchers to gain a better understanding of the brain regions involved in processing amplitude modulation throughout the auditory system, in order to go beyond correlational neuroimaging evidence and make causal links between brain and behaviour, methods that directly modulate neural systems are required. Transcranial magnetic stimulation (TMS) is a non-invasive brain stimulation technique that uses a time-varying magnetic field to induce an electrical current in targeted cortical tissue (Barker et al. 1985). This suprathreshold current discharges affected neurons and temporarily interferes with ongoing cortical activity (Klomjai et al. 2015). When combined with behavioural cognitive tasks, TMS can be a useful, reversible and safe method to probe the function of specifically targeted brain regions or networks (Hartwigsen 2015).

It has long been known that left and right auditory areas are structurally asymmetrical (Geschwind and Levitsky 1968). Abdul-Kareem and Sluming (2008) reviewed the evidence for structural variation of Heschl's gyrus (HG) and its included primary auditory cortex (PAC), and found large variability between individuals and hemispheres. Most studies reported evidence for a leftward bias for increased HG volume, surface area and length and this hemispheric structural asymmetry was further linked to increased language learning abilities. This leftward structural asymmetry is also found in the planum temporale (PT), a secondary auditory area known to be involved in AM processing at low rates (Giraud et al. 2000; Hutsler 2003). The structural connectivity of auditory areas has also been shown to be asymmetrical (Mišić et al. 2018). The right hemisphere has been shown to be more widely connected to multiple brain regions, despite the average path length being the same in the left and right hemispheres. The primary driver of this measured connectivity asymmetry is the abundance of interhemispheric connections from the right auditory areas.

Evidence for functional asymmetries has been reported from studies using neuroimaging and neurostimulation methods (Morillon et al. 2010; Bueti et al. 2008) but the precise nature of these asymmetries and their functional relevance to speech processing remains controversial. One of the most prominent hypotheses arising from these studies is that left hemisphere areas are specialised for processing temporal structure whilst right hemisphere areas are more sensitive to spectral cues (Zatorre and Belin 2001; Zatorre et al. 2002). A separate account, the Asymmetric Sampling in Time hypothesis (Poeppel 2003), proposes that left hemisphere auditory areas integrate over shorter temporal windows (∼25 ms) whereas right hemisphere auditory areas integrate over longer temporal windows (∼250 ms). The length of these integration windows align with the syllabic (∼4 Hz) and phonemic (∼40 Hz) rates of speech and closely correspond to key oscillatory frequency bands (theta and low gamma, respectively), suggesting functional importance. Boemio et al. (2005) provided support for this account when they reported fMRI responses to auditory stimuli that were parametrically varied in temporal content and found evidence for two distinct timescales of processing. These two models are not mutually exclusive and some aspects are interrelated, for example, a longer integration window (or lower sampling rate) would enable right hemisphere auditory areas to process spectral information in speech that usually occurs over a longer timescale (e.g. prosodic cues). In contrast, a higher sampling rate would allow better sensitivity to temporal information in the left hemisphere. There is mounting neurophysiological evidence for an overarching system specialised for processing speech at multiple timescales, (Giraud and Poeppel 2012; Poeppel 2014; Ding et al. 2016) likely involving distinct frequencies of oscillatory activity in the theta and low-gamma ranges (Ghitza 2011; Gross et al. 2013), with tracking of the acoustic speech envelope (Kubanek et al. 2013) and oscillatory phase resetting driven by sharp transients likely playing a role in synchronisation (Doelling et al. 2014).

The existence of asymmetric temporal sampling remains contentious and there is no current consensus about why these hemispheric asymmetries exist. While the asymmetric models of auditory processing may provide a useful framework and can account for some aspects of auditory processing, these simplified dichotomies have been criticised, primarily for mischaracterising the true nature of the speech signal (McGettigan and Scott 2012). Indeed, the appealing simplicity of these domain general models may have led to a true asymmetry being overlooked with right hemisphere auditory areas consistently showing a preference for slowly changing or longer sounds and the left hemisphere homologue not consistently showing specialisms for rapid temporal content (Scott and McGettigan 2013). Testing these asymmetric models using novel tools is one way to gain a more detailed understanding of how the auditory system processes amplitude modulation. Combining online TMS with AM detection is a promising approach as TMS has been shown to be effective at modulating many auditory and language-related abilities (Hartwigsen 2015), including similar low-level aspects such as duration discrimination (Bueti et al. 2008). Further, the areas involved in processing AM can be reliably located using fMRI and these can be used as TMS targets to ensure an area responsible for AM processing is targeted in each participant.

The current study used online double-pulse TMS to further investigate how AM is detected by the human auditory system. Behavioural AM detection tasks were first completed to identify the detection threshold for each participant. Cortical areas that are involved in processing AM were then located using a sparse fMRI paradigm. TMS was delivered to these left or right auditory cortex targets while participants performed an AM detection task at a modulation rate of 4 or 40 Hz. This online 'virtual lesion' TMS approach is a common and reliable technique which allows researchers to temporarily simulate naturally occurring lesions (such as those found in stroke) in neurotypical participants. AM rates of 4 and 40 Hz were chosen to align with the syllabic and phonemic rates of speech, while a broadband noise carrier was used in order to remove higher order aspects of speech, such as semantic content, while retaining a similar frequency bandwidth.

This study was designed to answer two specific research questions:

- i) Can low-level auditory processing (AM detection) ability be affected by online TMS to left and/or right auditory cortex?
- ii) If so, do these effects differ between the left and right auditory areas as implied by lateralised accounts of auditory processing.

Two pre-registered hypotheses were formulated to address these research questions (Partridge 2017):

- i) It was predicted that AM detection performance would differ between conditions where participants received real TMS and sham TMS.
- ii) It was predicted that these effects would show an interaction with the hemisphere of stimulation such that left hemisphere stimulation will have a greater effect on 40 Hz AM performance and right hemisphere stimulation will have a greater effect on 4Hz AM performance.

3.3 Materials and Methods

3.3.1 Subjects

Ten subjects (7 male; mean age = 21.1 years, SD = 3.76) participated in the experiment. Subjects responded to an advert and were paid or reimbursed in course credit for participation except for one participant who was a member of the research team. All subjects were TMS and MRI eligible, right-handed, native English speaking and had self-reported normal hearing. The project was approved by the Research Governance Committee, York Neuroimaging Centre, University of York and conformed to the guidelines of the Declaration of Helsinki. All participants gave written informed consent. One participant withdrew from the experiment, leaving a total of 9 participants.

3.3.2 Experimental Design

Audiogram

Subjects underwent an air conductance pure tone audiogram to determine that hearing thresholds were above eligibility requirements. Audiograms were conducted using an Interacoustics AD226 audiometer (Interacoustics, Middelfart, Denmark) and Telephonics TDH-39P earphones (Telephonics, NY, USA). Subjects were excluded from further testing if they did not meet the threshold criteria of $\leq$20 dB HL at octave frequencies from 250 – 8000 Hz in both ears.

Psychophysical Testing

A double-walled, sound-attenuated booth was used for psychophysical testing. Auditory stimuli were presented diotically via magnetically shielded Etymotic ER-2 headphones (Etymotic Research Inc., IL, USA). A modified version of the two-alternative forced choice (2AFC) task used by Viemeister (1979) was used to determine psychophysical thresholds for AM detection. Each stimulus consisted of five continuous 500 ms periods of broadband noise (low-pass filtered at 8 kHz). In each trial, either the second or fourth period was chosen at random to be sinusoidally amplitude modulated. The first, third (centre), and fifth 500ms period were always unmodulated. This stimulus forms one continuous 2.5 s percept. LEDs were used to indicate the two test intervals and feedback was given after each trial. Participants indicated which interval (either interval two or four) contained modulation by means of a key press. Participants were tested separately at AM rates of 4 and 40 Hz with 20 practice trials delivered at each rate before testing began. The starting phase of the modulation was chosen randomly from a uniform distribution ranging from 0 to 2 π and the standard and comparison interval power were normalised to match the RMS of the centre period. The AM depth of the modulated interval was systematically modified between trials using an adaptive staircase procedure (Levitt 1971). Two interleaved staircases were used (1-up, 2-down; 1-up, 3-down). Before each trial, a random choice was made between the remaining staircases and this continued until each staircase had reached 14 reversals. A step size in modulation depth of 1 dB was used for the first 4 reversals of each staircase, which decreased to 0.5 dB for the next 4 reversals and to 0.25 dB for the final 6 reversals. Participants completed three runs (six staircases in total) at each modulation rate. For 4 and 40 Hz separately, an estimate of the 70% threshold from each run was calculated from the respective psychometric function on a per-subject basis. The largest of the three modulation depth thresholds was discarded and the mean of the other two was used as the threshold depth for that subject in the TMS experiment. All stimuli were generated at a sample rate of 44.1 kHz using Python and presented using PsychoPy (Peirce 2007) and the PyAudio module (`https://people.csail.mit.edu/hubert/pyaudio/`) through an EMU 0204 24-bit DAC (Creative Technology Ltd, Singapore). Stimuli were presented at 70 dB SPL as calibrated using an artificial ear (Bruel & Kjaer, Naerum, Denmark; 4153) and a sound level meter (Bruel & Kjaer, Naerum, Denmark; 2260).

fMRI Acquisition and Analysis

Subjects wore earplugs underneath the sound-attenuating headphones from the fMRI-compatible auditory stimulus delivery system (MR Confon, MR Confon GmBH). fMRI data were acquired using an ISSS fMRI sequence (Schwarzbauer et al. 2006). The sound level of the scanner noise, not accounting for attenuation provided by earplugs and ear defenders, was 81 dB SPL during the quiet period and 98 dB SPL during the acquisition period. During the acquisition periods, whole head fMRI data (GE-EPI, TR = 2 s, TE = minimum full, FA = 90°) were collected using a GE Signa HDx 3T system (General Electric, WI, USA). A 128×128 pixel matrix with a field of view of 25.6 cm was used, giving an in-plane resolution of 2×2 mm. 24 interleaved slices were collected with a slice thickness of 2 mm. Slices were oriented parallel with the Sylvian fissure while ensuring bilateral coverage of Heschl's gyrus and planum temporale.

Stimulus conditions were silence, noise, 4 Hz AM noise, and 40 Hz AM noise. All noise carriers were broadband (low-pass filtered at 8 kHz) and the AM stimuli were fully modulated ($m = 1$). Stimuli were presented for 6 seconds followed by an acquisition period of 8 seconds. Each stimulus was presented 6 times in each fMRI run and three runs were performed for each subject, each run lasting approximately 6 minutes.

Data were analysed using Feat 5.98, part of FSL-4.1 (Smith et al. 2004) along with custom scripts which implemented filtering of the temporally non-contiguous data. Full details of the analysis procedures can be found in Hymers et al. (2015), with the exception that in this study, spatial smoothing was performed using a kernel with a FWHM of 2 mm. Data from the multiple fMRI runs for each subject were combined using a fixed-effects analysis. A contrast of 4Hz and 40Hz AM noise combined over unmodulated noise was performed. The results were corrected for multiple comparisons using a cluster thresholding procedure ($Z > 2.3$, $p = 0.05$) (Worsley 2001). TMS targets were defined individually for each subject using a superficial maxima voxel in a significant cluster on posterior superior temporal gyrus (pSTG) / planum temporale (PT) for each hemisphere.

Sagittal whole head structural T1-weighted data (3D FSPGR, TR = 7.8 ms, TE = minimum full, FA = 20°, Matrix 256 × 256, FOV = 29.0 cm, 176 slices, Slice thickness = 1 mm) were collected in the same scanning session as the fMRI data for participants who did not already have an available structural MRI.

TMS

A 2x3 repeated measures design was used during the main TMS experiment. Experimental factors were AM rate (2 levels; 4 Hz, 40 Hz) and TMS (3 levels; sham, left, right). The dependent variable was AM detection accuracy. All subjects completed six TMS sessions in total; one for each permutation of AM rate and TMS condition. The order of sessions was approximately counterbalanced across subjects using a latin square design. TMS sessions were separated by at least 24 hours. A Brainsight frameless stereotaxic system (Rogue Research, Montreal, Canada) with chin rest and forehead support was used to monitor coil placement and ensure accurate TMS delivery. A Magstim Super Rapid^2 stimulator (Magstim, Whitland, UK) and three identical 80mm figure-of-eight TMS coils (Magstim, Whitland, UK; 80mm diameter of each coil wing) were used for pulse delivery. TMS coils were always oriented with the handle pointing horizontally in an anterior direction, parallel with the midline. Stimulation consisted of a double-pulse protocol with pulses separated by 50 ms. The first TMS pulse in each pair was delivered 35ms post interval onset and the second pulse in the pair was delivered at 85ms post stimulus onset. TMS pulse intensity was set to 70% of the maximum stimulator output (maximum = 2.6 T). This intensity was chosen as compromise between the increasing effectiveness of stimulation and decreasing tolerability of TMS at higher intensities. Before the first experimental TMS session, test pulses were delivered to ensure that subjects could tolerate this TMS protocol. One TMS target in the left hemisphere and one in the right hemisphere were obtained using the fMRI data as previously described. Sham TMS was performed by placing the coil at the vertex and rotating by 90°to reduce any magnetic stimulation to a negligible level but, crucially, retain peripheral factors such as the auditory artefact and pressure placed on the head due to the coil (Lisanby et al. 2001). Audio was delivered during the experiment using the same Etymotic ER-2 headphones and earpieces as were used during psychophysical testing.

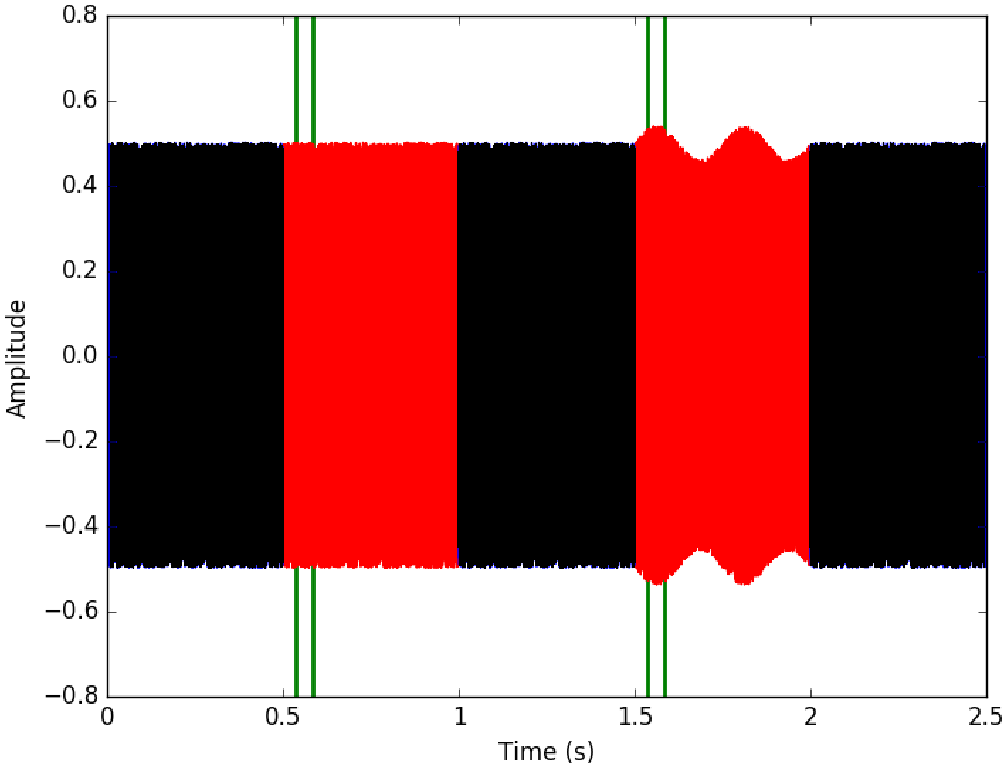

Figure 3.1: An example trial from the main task run. During the main task runs, the AM depth of the target interval remained constant (set per participant at their 70% threshold value). The AM rate (4 or 40 Hz) and TMS target (left, right or sham) did not differ within a session.

The design of the 2AFC task was identical to that used during baseline psychophysical testing, apart from the AM depth of the target interval was fixed (at 70% threshold), the TMS was delivered and participant feedback was removed. This threshold percentage was chosen to allow increases and decreases in AM detection ability to be measured. Two double-pulse TMS bursts were delivered within each trial. One double-pulse burst was always delivered at 35 ms post-onset of the second noise period and another double-pulse burst was always delivered 35 ms post-onset of the fourth noise period. Three blocks of 50 trials were performed, with short breaks between each block when the TMS coil was replaced to prevent overheating. A total of 600 TMS pulses were delivered in each session, at frequencies within the established safety guidelines (Rossi et al. 2009).

Statistical Analysis

A generalised linear mixed-effects model (GLMM) was fit to the AM detection accuracy data. This technique accounts for interdependence of the data arising from repeated measurements from the same participants; it also allows adjustments for non-normal distributions (Baayen et al. 2008; Dixon 2008; Jaeger 2008). A logit link function was used to account for the binomial distribution of the raw data. The model included fixed effects of TMS condition and AM rate and a TMS condition by AM rate interaction term. Participant was included as a random effect. Random intercepts and slopes for TMS condition and AM rate were estimated for each participant. The random effects structure was the maximal justified by the design (Barr et al. 2013); the model with maximal random effect structure was fit then systematically reduced until convergence. Treatment coding was used to measure the difference from the baseline condition of sham TMS with 4Hz AM to the other conditions. Finally, parametric bootstrapping was conducted on the model parameters to obtain 95% confidence intervals. This model design was chosen as it is sufficient to test the experimental hypotheses without including excess parameters that increase the risk of overfitting. All modelling was conducted using the *lme4* package (Bates et al. 2015) in *R* software (R Core Team 2017). Parametric bootstrapping was conducted using the *pbkrtest* package (Halekoh and Hojsgaard 2014).

3.4 Results

3.4.1 Psychophysics

The modulation depth at which 70% accuracy was achieved was calculated individually for each participant, for each AM rate (4 and 40 Hz). The mean threshold modulation depth across participants for the 4 Hz condition was -23.3 dB and for the 40 Hz condition was -21.5 dB. These values correspond to 0.069 m and 0.085 m respectively and are consistent with previously reported TMTFs (Viemeister 1979). A paired t-test showed a statistically significant difference between thresholds for 4 Hz and 40 Hz AM detection ($t(8) = -4.17$, $p = 0.003$). This suggests that participants had consistently higher thresholds for detecting 40 Hz than for detecting 4 Hz and their threshold values.

Table 5.1 shows the 70% threshold values obtained from the behavioural runs for the AM detection task at 4 and 40 Hz for each participant. These depth values were used for each trial in the TMS sessions.

Participant	4 Hz threshold (dB)	40 Hz threshold (dB)
P1	-21.7	-20.7
P2	-24.5	-23.6
P3	-23.9	-21.9
P4	-22.4	-22.3
P5	-23.3	-20.5
P6	-26.0	-21.8
P7	-21.9	-20.4
P8	-21.9	-19.4
P9	-24.0	-23.4

Table 3.1: AM detection 70% threshold values in dB for each participant. These values were used in the target interval in the TMS runs. dB values are computed as $20\log(m)$.

3.4.2 fMRI derivation of TMS Targets

TMS targets in MNI-152 co-ordinate space for all participants can be found in Table 3.2.

Participant	Left hemisphere target	Right hemisphere target
1	-58, -14, 0	52, -20, 4
2	-60, -26, 10	52, -20, 4
3	-48, -26, 4	66, -14, 10
4	-56, -16, 4	56, -20, 6
5	-44, -32, 10	64, -26, 16
6	-58, -38, 10	62, -14, 6
7	-66, -30, 10	64, -26, 14
8	-60, -16, 8	60, -32, 18
9	-64, -40, 14	62, -24, 12
Mean	-57.1, -26.4, 7.8	59.8, -21.8, 10

Table 3.2: MNI-152 co-ordinates for TMS targets in left and right hemispheres for all participants. The average co-ordinate across all participants is also given.

Figure 5.2 shows targets in both hemispheres for an example participant.

TMS target distance measurements were recorded for each TMS pulse. This is a measure of the distance from the centre of the TMS coil to the fMRI-defined target. One session lacked these measurements due to a technical issue. A linear mixed effects model was fit to this data, with a fixed effect of TMS hemisphere and a random effect of participant. Parametric bootstrapping was conducted on the model parameters to obtain 95% confidence intervals. There was a negligible mean difference in target distance of -0.228 mm, 95% CI [-6.99, 6.59] between left and right hemisphere conditions. There was no overall correlation between the AM detection accuracy scores and the target distance measurements (r = -0.065, p = 0.949). This was calculated by correlating the average distance in each session with the average accuracy in that session.

TMS positioning error measurements were also recorded for each TMS pulse. This is a measure of the distance between the fMRI-defined target and the estimated pulse trajectory. One session lacked these measurements due to a technical issue. A linear mixed effects model was fit to this data, with a fixed effect of TMS hemisphere and a random effect of participant. Parametric bootstrapping was conducted on the model parameters to obtain 95% confidence intervals. There was a negligible mean difference in positioning error of 0.019 mm, 95%

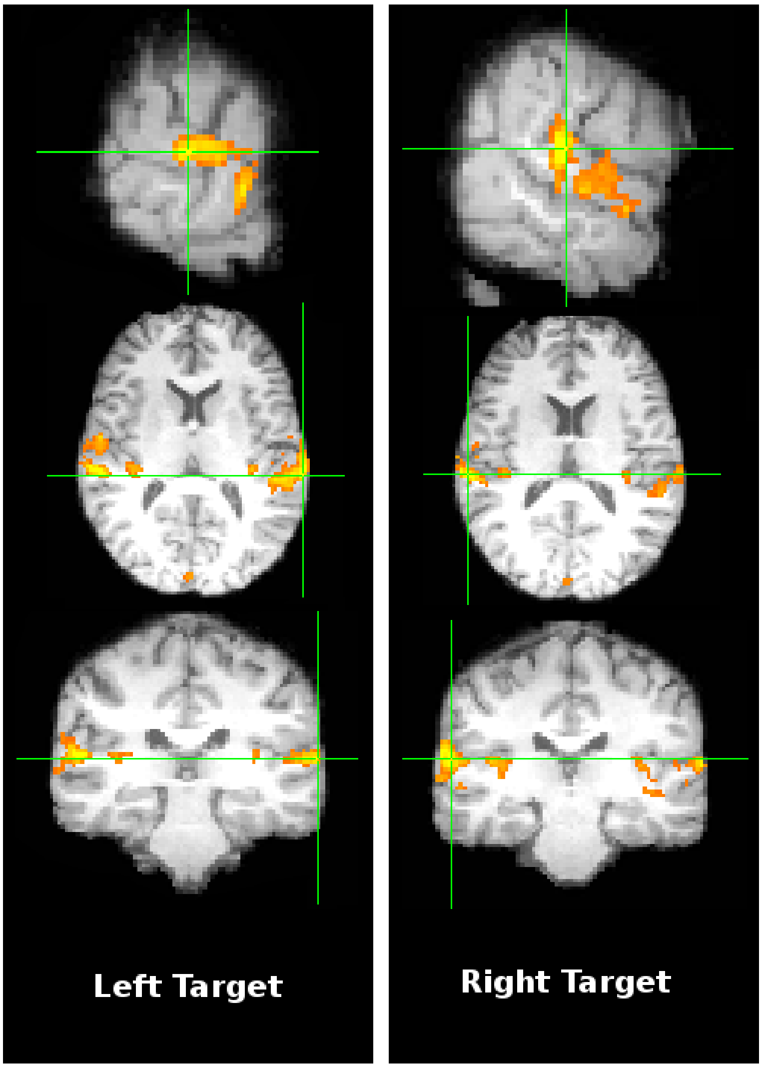

Figure 3.2: fMRI TMS target localisation data from an example participant. Overlay shows clusters of activation (Z >2.3; p <0.05) in bilateral pSTG/PT in response to a combination of 4 and 40 Hz sinusoidally amplitude modulated broadband noise conditions (combined) over unmodulated noise. TMS targets are indicated by the crosshairs. Clusters are shown overlaid on MNI-152 data - targets are back transformed into individual co-ordinates. Target locations for this participant in MNI-152 space are: Left: $-66, -30, 10$ and Right: $64, -26, 14$

CI [-0.038, 0.078] between left hemisphere and right hemisphere conditions. The overall mean error was 0.273 mm, 95% CI [0.202, 0.324].

3.4.3 TMS Behavioural Analysis

AM Rate (Hz)	Sham TMS	Left TMS	Right TMS
4	69.0	65.8	63.6
40	55.1	54.4	55.3

Table 3.3: Mean AM detection accuracy for all conditions.

The mean AM detection accuracy for each condition can be found in Table 3.3. These results show that on average, accuracy in all 40 Hz conditions was much lower than in 4 Hz conditions. The effects of extraneous factors related to TMS delivery, such as the auditory artefact and behavioural arousal effects were assessed using the sham TMS conditions. As these factors were present in both real and sham TMS conditions, using sham TMS scores as baselines for comparison incorporates this control. As AM depth in the main sessions was calibrated to the 70% threshold AM depth value for each participant, it was expected that the average accuracy in the sham condition would approximate this. The mean accuracy of 69.0% in the 4 Hz AM detection task with sham TMS suggests that the AM depth calibration performed well in this condition, however the mean accuracy of 55.1% in the 40 Hz AM detection task with sham TMS is near chance level (50% for a 2AFC task). This has a large impact on the results for all 40 Hz conditions as there is little scope for TMS to decrease accuracy due to a floor effect.

Accounting for sham performance for each participant, the main results can be seen in Figure 5.4.

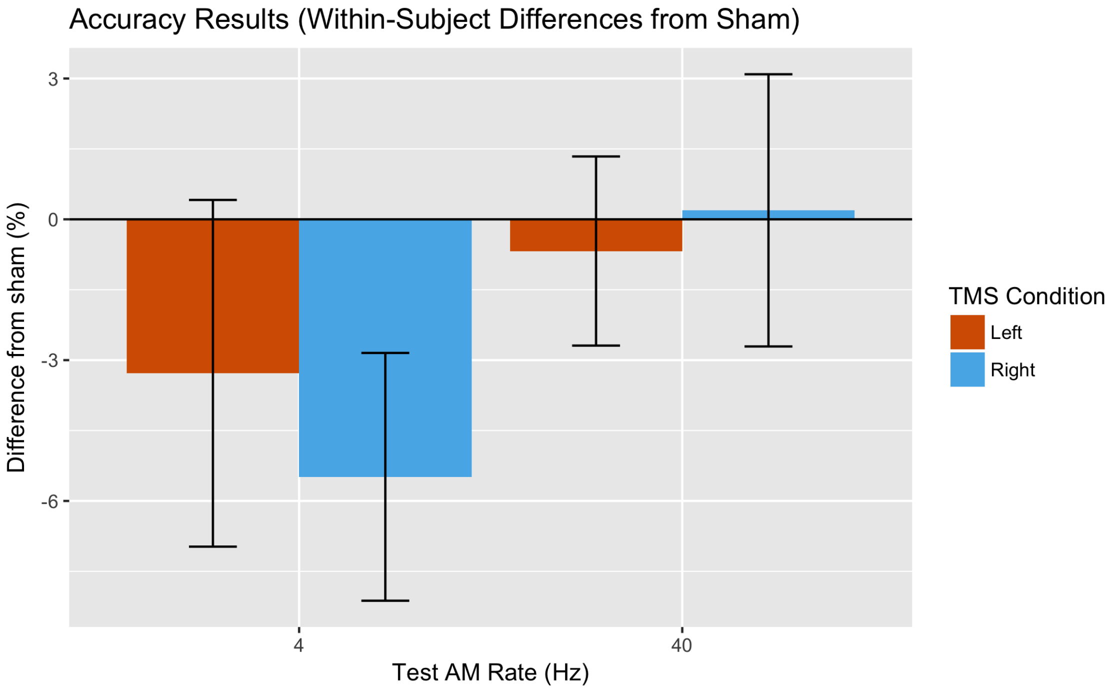

Figure 3.3: Differences in AM detection accuracy (percentage correct) between the sham TMS condition and left or right TMS conditions for each AM rate; calculated as the mean of within-subject differences for each participant. Error bars show the 95% confidence interval.

Main GLMM Analysis

	Dependent variable:		
	Accuracy	95% CI	
Baseline	0.834	0.525	1.137
Left Hemisphere TMS	−0.157	−0.421	0.112
Right Hemisphere TMS	−0.241	−0.510	0.026
AM Rate	−0.627	−0.930	−0.321
Left Hemisphere TMS x AM Rate	0.128	−0.101	0.348
Right Hemisphere TMS x AM Rate	0.249	0.021	0.474

Table 3.4: Main GLMM analysis results. Accuracy parameters and 95% CIs are in logit units. 'Baseline' represents the baseline condition of sham TMS with 4 Hz AM. All other coefficients represent the difference from this baseline.

Factors included in the final model were TMS condition (3 levels: sham, left hemisphere and right hemisphere) and AM rate (2 levels: 4 Hz and 40 Hz). Model fit values were as follows: AIC = 439.2, BIC = 471.0. The 4 Hz AM rate with sham TMS condition was selected as the baseline condition so all coefficients represent the change from this condition. The coefficient for the baseline condition represents the overall mean intercept and the other coefficients represent the (population-level) difference from this baseline for each condition. This type of treatment contrast allows the sham control to be taken into account for every comparison. The coefficients are presented in logits as this preserves the magnitude (effect size) and direction of the effect. Further, this aids interpretation of significance as confidence intervals are symmetrical about the point estimate. Conventional measures of significance like p-values are not possible to calculate with GLMMs as the required denominator degrees of freedom value is unknown (which is required to calculate the F-statistic). However, parametric bootstrapping can be used to calculate confidence intervals which can indicate approximate significance. 95% confidence intervals have been calculated here, so if 0 is not included between the lower and upper limits of the confidence interval, that parameter is approximately significant at the 0.05 alpha level.

There was a clear overall difference in accuracy dependent on AM rate, as shown by the negative *AM Rate* coefficient. However, the positive *Right Hemisphere TMS x AM Rate* coefficient shows that there is an interaction between hemisphere of stimulation and AM rate such that TMS to the right hemisphere affected 4 Hz AM detection accuracy to a greater degree than it affected 40 Hz AM detection (accounting for sham).

To explore these effects further, the raw data was plotted for each individual. This can be seen in Figure 3.4. It is notable that two clear response patterns are visible and this may represent separate subgroups of responders and non-responders. Generally, in the 4 Hz condition, if a participant shows a disruption in performance with left hemisphere stimulation, they also show a disruption with right hemisphere stimulation and vice versa (apart from two participants who show the opposite effect for left and right hemisphere stimulation. In the 40 Hz condition, a similar pattern can be seen, although less pronounced due to the smaller range of scores. High variability in the direction of brain stimulation effects is often found and this type of response distribution is common (López-Alonso et al. 2014).

3.5 Discussion and Conclusions

This study was designed to investigate whether online TMS to human auditory cortex could modulate the processing of amplitude modulation. The study also aimed to further investigate the lateralisation of AM processing by comparison with existing asymmetric models of auditory function. Two pre-registered hypotheses were tested (Partridge 2017).

The first hypothesis predicted that active TMS to left or right auditory areas would affect AM detection accuracy when compared to sham TMS. A decrease in 4 Hz AM detection accuracy was shown only when TMS was delivered to the right hemisphere. This novel causal evidence shows that the right posterior superior temporal gyrus (pSTG) region is crucial for processing 4 Hz AM. This finding also demonstrates the utility of TMS for modulating AM processing and

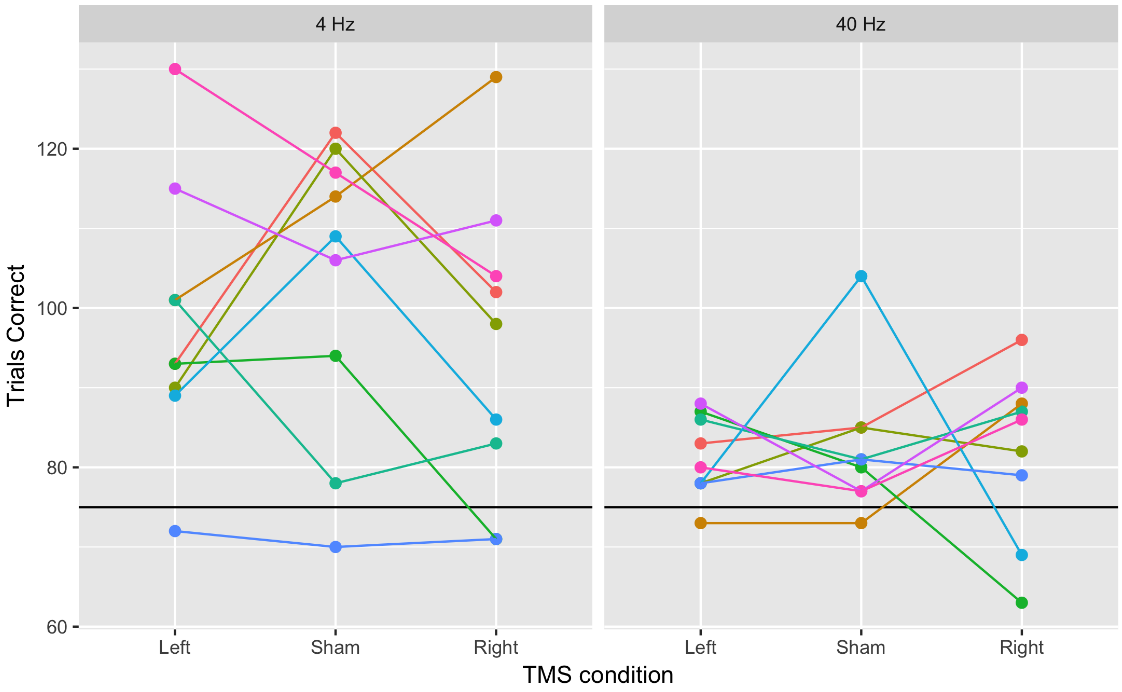

Figure 3.4: Raw accuracy scores for each participant and each condition. Colour is used to represent data from each different participant. The black horizontal line represents chance performance. Connecting lines are used to help visually identify which scores are from which participant.

for further investigating the processing of low-level speech-relevant stimuli.

The second hypothesis predicted that left hemisphere stimulation would have a greater effect on 40 Hz AM performance and right hemisphere stimulation would have a greater effect on 4Hz AM performance. These asymmetric predictions were in line with the Asymmetric Sampling in Time hypothesis (Poeppel 2003). The disruptive effect of TMS to the right, but not left hemisphere and specifically affecting 4 Hz AM detection, is in accordance with a specialisation of the right auditory areas for temporal processing on longer timescales. No significant differences were found in the 40 Hz AM rate condition, however the sham TMS condition provides evidence that this is due to a floor effect. This outcome leads to a conclusion that the data collected from the 40 Hz AM conditions provides no support for or against asymmetric processing of this AM rate.

3.5.1 Effect of TMS on 4 Hz AM Detection

The main result from this study was a significant decrease in 4 Hz AM detection accuracy when TMS was delivered to the right auditory cortex during presentation of AM noise. This effect demonstrates that there is a causal link between right pSTG/PT and AM detection. Whilst the exact mechanism underlying online TMS effects is still unknown, it is suspected that behavioural effects may specifically result from an early suppression of neural firing by suprathreshold TMS (Moliadze et al. 2003; Silvanto and Cattaneo 2017). As the TMS targets were located on an individual basis, it is important to note that this effect may not generalise to all locations on the pSTG or PT, but just to the areas located using this specific functional localiser and contrast within each individual. It is known that PT is a computational hub that is involved in the extraction of complex modulation (Griffiths and Warren 2002) and this finding suggests that pSTG/PT may be the neural locus of (at least part) of the proposed modulation filterbank system (Dau et al. 1997a; Jepsen et al. 2008). Whilst the simple 'virtual lesion' disruptive explanation accounts for the TMS effect observed here, it is important to consider what similar investigations have found while concurrently measuring the neural signals, typically with electroencephalograpy (EEG). Related investigations have shown that auditory areas entrain to rhythmic sensory stimulation and this entrainment is likely to play a functional role (Zoefel et al. 2018). TMS delivered during neural oscillatory activity has been shown to reset the phase of these oscillations, so it is possible that TMS has disrupted this entrainment process and subsequently, AM detection has been affected (Van Der Werf and Paus 2006). However, it is unclear if entrained oscillatory activity would emerge with the short periods of AM used here, particularly as TMS was delivered early (35ms) in the AM interval. Further, rhythmic TMS has been shown to actively induce oscillations both at specific TMS burst frequencies (e.g. 5-pulse burst at alpha frequency: Thut et al. 2011), or at the intrinsic resonance frequencies of specific networks with a single TMS pulse (Rosanova et al. 2009). As no direct measures of neural activity were made in this study, it is not possible to demonstrate the existence of entrainment or any effect TMS may have had on intrinsic or induced oscillations. So, while this is an interesting novel result which future studies can expand on, the explanation for TMS effects that this behavioural measure affords is perhaps unsatisfying. Combining direct neural measures with TMS would be required to understand this effect at the neural level and this is a promising future direction, based on successfully applied information-based approaches (Romei et al. 2016). In particular, using EEG to measure the effects of single pulse TMS or rhythmic TMS at specific frequencies would be interesting to explore further and would help to elucidate the neural basis of this effect.

There was no significant effect of TMS to the left auditory cortex on 4 Hz AM detection accuracy. In combination, these findings provide support for a right hemisphere specialisation for processing over longer timescales, but they are not sufficient to fully support asymmetric accounts, where claims are made for both hemispheres, such as the Asymmetric Sampling in Time hypothesis (Poeppel 2003). This finding aligns with the observations of McGettigan and Scott (2012), that the right hemisphere may be specialised whereas the left may not, but care must be taken when this dissociation across hemispheres may be due to methodological factors. fMRI studies have consistently shown that both left and right pSTG/PT areas are activated by low frequency AM, suggesting some involvement of the left hemisphere (Giraud et al. 2000; Harms et al. 2005), but this study provides initial evidence that the left pSTG/PT may not be crucial for processing 4 Hz AM. However, this conclusion is tentative and is accompanied by some important caveats. Generally, interpreting the lack of an effect when it is not part of a clear dissociation is difficult as it could be the result of a number of factors. Further, a small (non-significant) decrease in 4 Hz AM detection accuracy was also found with TMS to the left hemisphere, so it could be that the current study lacks the statistical power to detect this smaller effect suggested by the fMRI evidence. Also, there is variability in the effects of different TMS protocols and this effect of this specific protocol does not necessarily generalise to other TMS protocols. It is possible that the targeted area in the left hemisphere is less susceptible to disruption by this specific TMS protocol compared to the right hemisphere homologue.

One promising explanation for this asymmetrical effect of TMS arises from the asymmetrical connectivity profiles of these areas. It is known that auditory areas in the left and right hemisphere are structurally and functionally asymmetrical (Mišić et al. 2018; Boemio et al. 2005). This asymmetry is also reflected in the connectivity profiles of these areas; the right hemisphere is more widely connected, particularly with regards to interhemispheric connections (Mišić et al. 2018). Further, the effects of TMS have been shown to be asymmetrical, with right hemisphere stimulation increasing neural activity in the contralateral left hemisphere auditory regions whereas no contralateral activation was found when the left hemisphere was directly stimulated (Andoh and Zatorre 2013). It has since been shown that this asymmetrical TMS effect could be related to asymmetrical functional connectivity (Andoh et al. 2015). These studies used a repetitive TMS protocol that differs from the protocol used here, however, it may explain the hemispheric dissociation seen in the 4 Hz AM detection results. Clearly, the simple assumption of a direct mapping of location to behavioural effects is not as straightforward in auditory areas as may be expected in motor areas, where most TMS methods are initially developed. Further exploration of the functional connectivity of these auditory areas is required and investigations with different TMS protocols are necessary to conclude that the hemispheric dissociation found here is due to differences in connectivity.

3.5.2 No Effect of TMS on 40 Hz AM Detection

No significant effect on 40 Hz AM detection sensitivity was measured with TMS delivered to either hemisphere. This is likely due to the average performance being too close to chance level, leading to a floor effect for decreases in accuracy due to TMS. Although interpretation is difficult, a recently proposed framework that accounts for both 'virtual lesion' and state-dependent TMS effect may be applied in order to use this condition to inform future studies (Silvanto and Cattaneo 2017; Silvanto et al. 2018). Following the framework posited by Silvanto and Cattaneo (2017), if left hemisphere neural populations are tuned to 40 Hz AM, a suprathreshold pulse in their suppressive range would inhibit performance. This aligns with the prediction made for 40 Hz neurons with left hemisphere stimulation and cannot be observed in these data due to the floor effect. However, the framework posits that neurons not tuned to the stimulus could be in a facilitatory or inhibitory range, depending on the intensity of the stimulation and this results in clear testable predictions for future studies. It predicts that, if these neurons are not tuned to 40 Hz AM, then as stimulation intensity decreases, a point will be reached where a behavioural facilitation is observed in this condition. As this facilitation would manifest as an increase in accuracy, this could be observed even in the case of near-chance baseline accuracy. At this point, it is predicted that neurons tuned to 40 Hz would be in their inhibitory range, whereas neurons not tuned to 40 Hz (e.g. tuned to 4 Hz), would be in their facilitatory range. If there are overlapping populations within the same hemisphere responsible for processing 4 and 40 Hz AM, this would result in a behavioural dissociation where 40 Hz AM detection is disrupted, but 4 Hz AM detection is facilitated. Showing this dissociation within the same target and with the same intensity of TMS would be a powerful demonstration and would more strongly support the lack of specialism in the left hemisphere for particular AM rates. As a suprathreshold TMS intensity was used here and no facilitation of performance was observed, it can be assumed that neurons were in their inhibitory range, so future investigations could use this intensity as a starting point and decrease it from there. Further, the framework would also predict that as intensity decreased even further, it would reach a point where no effect was observed on neurons not tuned to 40 Hz AM (e.g. tuned to 4 Hz AM), but a facilitatory effect was observed on neurons tuned to 40 Hz AM. However, this framework was primarily developed using the results of visual studies and the extent that it holds for auditory areas is not clear, particularly when the aforementioned asymmetry of connectivity is considered.

Another key aspect to consider for future studies is the target selection method. In order to avoid biasing the fMRI-defined TMS targets to populations responsible for processing either rate, they were chosen based on a contrast of 4 and 40 Hz AM noise combined over unmodulated noise. Indeed, when the depth of these targets were analysed, they were found to not differ significantly. However, previous research has found that areas involved in processing lower rates of AM tend to be located more laterally (Herdener et al. 2013; Harms and Melcher 2002; Harms et al. 2005). As the magnetic field generated by TMS originates laterally from outside the skull (and the field strength drops off dramatically with increased distance from the coil), TMS is more likely to affect cortical areas that are lateral to the target, as opposed to medial. This could result in lateral areas that are specialised for processing 4 Hz being inadvertently stimulated moreso than medial areas that are specialised for processing 40 Hz. When targeting a deeper target, on PT for example, the induced stimulation current may be more likely to affect 4 Hz selective populations than 40 Hz selective populations. It is yet to be shown if TMS can affect 40 Hz AM detection, so this is a key consideration for future studies. It is difficult to make conclusions about the true asymmetry of AM processing as less weight can be given to conclusions regarding the 40 Hz data, due to near-chance accuracy. However, these data can still be used to inform future research and highlight important considerations, particularly with regards to similar psychoacoustic paradigms.

3.5.3 Accuracy Discrepancy in the 40 Hz AM Conditions

Mean AM detection accuracy in the 40 Hz condition with sham TMS was close to chance level at 55.1%. This suggests that the AM depth calibration method did not consistently estimate participants' 70% threshold in practice. This discrepancy could be due to caveats related to the task or the threshold estimation procedure.

Task Caveats

During the behavioural sessions, interleaved adaptive staircases were used to estimate the threshold. This technique starts at an AM depth where participants can easily complete the task and becomes progressively more difficult. In contrast, during the TMS sessions, the task began at the threshold depth and this difference may have led to a discrepancy in performance. This staircase procedure may have led to lower estimated thresholds than would be expected if the task was to immediately start at the same estimated threshold depth. The initial trials may have acted as practice trials or provided a recent reference for the auditory system, improving AM sensitivity. Further, no feedback was given during the TMS sessions, potentially increasing the difficulty of the task as the participant had no opportunity to self-correct, should they lose track of the specific aspect they had to detect. The combination of the different presentation method and the lack of feedback may have contributed to the discrepancy between threshold estimates obtained in the behavioural sessions and accuracy in the sham TMS sessions. However, the task procedure was identical for both AM rates and this discrepancy was only observed in the 40Hz conditions. This suggests that something may be inherently different in the function fitting and threshold estimation procedure between the two AM rates.

Threshold Estimation

The slope parameter is an intuitive candidate as a steeper slope in the 40 Hz condition would cause a minor underestimate of the threshold to result in a large difference in accuracy at that fixed depth. Another possible candidate is the function fit itself. In this study, the full psychometric function was fit and the threshold depth estimated from it. 3 separate runs were completed at each rate and the two lowest threshold values were averaged to get the AM depth for the main task. The function with the highest threshold value was removed to reduce the influence of any functions that were skewed by early mistakes and to somewhat factor out practice effects. However, the lowest threshold does not necessarily indicate the best function fit. To address this, goodness-of-fit metrics could be used to decide which thresholds to keep. Another consideration for future research is the choice of step size. The sequence of step sizes can be fixed and specifically chosen so that the average of n last reversals approximates targeted threshold percentages (Levitt 1971). This method would circumnavigate the psychometric function fitting thus avoiding biases that may arise from it. Alternatively, advanced adaptive methods (Treutwein 1995), such as updated maximum likelihood (Shen and Richards 2012; Shen et al. 2015) could be implemented to more efficiently estimate the psychometric function. These methods use variable step sizes as they adaptively select the next depth value in a track to maximise the information gain. A benefit of these more efficient method is that participants would be less likely to be fatigued due to long runs and threshold estimates may be less variable. It is also possible to modify the experiment so that TMS is delivered during adaptive staircase runs and the resulting threshold is then used as the dependent variable. This option would not feasible with the current staircasing method but a more efficient alternative such as these may allow it. However, it is still to be determined whether the discrepancy would be observed using these methods, as it may arise due to the task caveats rather than the function fitting. A final, simpler, alternative is to use a larger threshold percentage in the main TMS runs to allow for some underestimation of the threshold depth in the 40 Hz condition. These alternative methods may be more accurate and the estimated threshold AM depths may be less prone to this discrepancy.

3.5.4 Future Directions

There is increasing evidence for the importance of oscillatory activity (Haegens and Zion Golumbic 2018) and new neurostimulation techniques can be used to interface with this in increasingly nuanced ways (Romei et al. 2016; Herring et al. 2015). As auditory processing generally, and AM processing specifically, is inherently linked with oscillatory activity, this is an interesting area that would benefit from future research. It is likely that TMS in this study was affecting ongoing oscillations, however pulse timing is key. It may be possible to shift the TMS pulse timing in or out phase in order to entrain oscillations or disrupt existing rhythms. The auditory steady state response (ASSR) is a non-invasive electrophysiological measure that characterises how the auditory system responds to continuous periods of modulated sound, with peak ASSR amplitudes found at around 40 Hz modulation rates (Ross et al. 2000). However this method involves presenting long periods of auditory stimulation and it is less well known how the auditory system responds to short periods of AM, like those presented in this study. There are many interesting questions surrounding what influence the TMS timing would have on processing short periods of AM, and whether rate or hemispheric differences would affect this. Finally, high interindividual variability is common in brain stimulation studies and finding sources of this variance is important for reducing it in future studies (López-Alonso et al. 2014). Further investigating the possible responder/non-responder pattern seen in the per-participant data and identifying factors underlying this differential susceptibility to TMS would be valuable for future investigations.

3.5.5 Conclusions

This study used online dual pulse TMS to investigate the role of left and right auditory cortex in processing 4 and 40 Hz amplitude modulations. A disruption of 4 Hz AM detection was found when TMS was delivered to the right hemisphere only, compared with sham TMS. This is novel causal evidence for the role of the right auditory cortex in processing 4 Hz AM. It was thus demonstrated that TMS can be a useful and specific technique for modulating low-level auditory processing. Results for the 40 Hz AM condition were affected by behavioural thresholding issues so conclusions about asymmetrical processing are difficult to support. With respect to this, important considerations for future psychoacoustic and fMRI-guided TMS studies were discussed.

Chapter 4

Amplitude Modulation Detection Ability can be Differentially Affected using State-Dependent Transcranial Magnetic Stimulation

4.1 Abstract

Temporal changes in the speech envelope are crucial for intelligibility, and humans are particularly sensitive to amplitude modulations (AM) at frequencies commensurate with the syllabic rate (4Hz). Some asymmetric models of auditory processing have attempted to explain observed functional asymmetries between left and right auditory cortices in terms of temporal integration windows of differing lengths. In this paper we use an fMRI-localised, state-dependent, dual-pulse TMS paradigm to investigate auditory processing as exemplified by detection of sinusoidal amplitude-modulation of a broadband noise carrier. The effect of left and right hemisphere TMS on accuracy levels for detection of such modulations at rates of 4 and 40Hz under conditions of behavioural adaptation were examined.

Effects of TMS on AM detection were found only in the 40Hz detection task when TMS was applied to the left hemisphere. Further, this effect showed a dissociation based on the modulation used during behavioural adaptation, with a facilitation shown when participants were behaviourally adapted to 4Hz AM whereas a disruption was shown when participants were behaviourally adapted to 40Hz AM, contrary to initial predictions based on state-dependent TMS theory. These results show for the first time that it is possible to differentially affect AM detection ability using TMS and further emphasise the importance of subtle timing factors in TMS studies. The novel use of this technique in the auditory domain has implications for further elucidating auditory processing mechanisms in health and disease.

4.2 Introduction

The structure of human speech is highly dynamic, consisting of continuous changes in both amplitude and frequency content. Slow temporal amplitude modulations (Rosen 1992), are known to be crucial to speech intelligibility (Drullman et al. 1994a,b). Indeed, intelligibility remains high even under conditions of severely degraded spectral content, provided that these amplitude modulations are preserved (Shannon et al. 1995).

Psychophysical work has shown that humans are particularly sensitive to detecting AM at rates commensurate with the speech envelope and that AM detection can be characterised as a low-pass (Viemeister 1979; Rees et al. 1986; Kohlrausch et al. 2000) or band-pass process (Edwards and Chang 2013). It has been proposed that AM is extracted by a bank of filters, each tuned to different modulation rates (Dau et al. 1997a,b; Jepsen et al. 2008; Simpson et al. 2013; Xiang et al. 2013). The exact anatomical locus of such a modulation filterbank is currently unclear, although it is known that progressively higher areas of the auditory pathway show peak responsivity for lower AM rates (Harms and Melcher 2002; Harms et al. 2005), behaviour which has been compared to a cascade of modulation filters (Giraud et al. 2000).

In addition to differential sensitivity from lower to higher auditory areas, several hypotheses have been proposed to account for observed asymmetries in auditory processing. One hypothesis holds that the left hemisphere preferentially processes temporal structure whilst the right hemisphere is more sensitive to spectral cues (Zatorre and Belin 2001; Zatorre et al. 2002) whilst an alternative ('Asymmetric Sampling in Time'; Poeppel 2003) proposes that the right hemisphere integrates over longer timescales (∼250ms) than the left hemisphere (∼25ms). The longer of these timescales (∼250ms) corresponds closely with the syllabic rate of speech (∼4Hz). The shorter timescale (∼25ms) corresponds to a frequency of 40Hz, which is around the phonemic rate of speech. These two accounts are not mutually exclusive, for example, frequency modulations in speech (such as prosody) tend to occur over longer timescales, therefore a preference for longer slower modulations would overlap with a preference for spectral components generally. Dichotomous accounts of asymmetric processing have, however, been challenged - both due to a of lack of empirical support for left hemisphere temporal selectivity as well as the fact that such a division does not account for the true structure of the speech signal (McGettigan and Scott 2012).

In the current study, TMS was used to further investigate the behavioural detection of sinusoidal amplitude modulation of broadband noise. The study was designed with specific reference to asymmetric models of auditory processing.

A state-dependent, dual-pulse, TMS method was used to perturb judgements in a sinusoidal AM detection task at modulation rates of 4Hz and 40Hz. When TMS is delivered to neuronal populations in different activation states, state-dependent differential effects have been observed (Silvanto et al. 2007; Silvanto and Pascual-Leone 2008; Cattaneo and Silvanto 2008). In this context, state-dependent TMS essentially refers to the fact that the impact on behaviour caused by the TMS pulse is altered by the nature of the processing being performed by the neuronal population at the time at which the pulse is administered. These effects include behavioural facilitations for detection of adapted stimuli, suggesting that the common conceptualisation of TMS as a purely disruptive tool is inadequate (Romei et al. 2016). An overarching framework has recently been proposed that can explain classic 'virtual lesion' inhibitory effects of TMS and state-dependent effects (Silvanto and Cattaneo 2017). Combining state-dependent TMS with behavioural adaptation is a promising method for investigating AM detection for three principal reasons. Firstly, the perceptual effects of selective adaptation to amplitude and frequency modulated sounds are well characterised (Kay and Matthews 1972; Regan and Tansley 1979; Tansley and Suffield 1983; Bacon and Grantham 1989; Houtgast 1989; Wojtczak and Viemeister 2005; Kleinschmidt and Jaeger 2016). Secondly, state-dependent TMS potentially allows the dissociation of cortical functioning in brain areas with overlapping populations responsible for processing different rates of AM, as may be expected based on previous localisation studies (Barton et al. 2012; Herdener et al. 2013). Finally, as stimulation is delivered immediately prior to the auditory test stimulus in the state-dependent TMS paradigm, potential auditory masking due to the TMS click artefact is avoided. These three principal reasons provided the rationale for using the state-dependent TMS paradigm. Whilst different forms of neurostimulation has been used to modulate aspects of auditory perception such as temporal discrimination (Bueti et al. 2008; Heimrath et al. 2014), melody discrimination (Andoh and Zatorre 2011), phoneme categorisation (Rufener et al. 2016), and tinnitus (Weisz et al. 2012), to our knowledge, TMS has not previously been used to investigate AM processing.

We proposed three research questions.

- Can AM detection ability be affected by TMS?
- Do these effects show state-dependency?
- Do these effects show hemispheric differences consistent with lateralised accounts of auditory processing?

Three pre-registered hypotheses were formulated (Partridge 2016).

- It was predicted that AM detection performance would differ between real and sham TMS conditions.

- Secondly, it was predicted that TMS effects would show state-dependency; i.e. performance would only be improved by TMS when the modulation rates of the adaptor and test stimuli were congruent.

- Finally, it was predicted that effects would show an interaction with hemisphere of stimulation such that left hemisphere stimulation would have a greater effect on 40Hz AM detection performance and right hemisphere stimulation would have a greater effect on 4Hz AM detection performance.

4.3 Material and Methods

4.3.1 Subjects

Ten subjects (4 female; M = 23.1, SD = 5.14) participated in the experiment. Subjects responded to an advert and were paid or reimbursed in course credit for participation except for one participant who was a member of the research team. All subjects were TMS and MRI eligible, right-handed, native English speaking and had self-reported normal hearing. The project was approved by the Research Governance Committee, York Neuroimaging Centre, University of York and conformed to the guidelines of the Declaration of Helsinki. All participants gave written informed consent. One participant withdrew from the experiment and data from one participant had to be discarded due to acoustic noise caused by building work during the testing sessions, leaving a total of 8 participants. All data collection was undertaken after the pre-registration document (Partridge 2016) was made publicly available.

4.3.2 Experimental Design

Audiogram

Subjects underwent an air conductance pure tone audiogram to determine that hearing thresholds were above eligibility requirements. Audiograms were conducted using an Interacoustics AD226 audiometer (Interacoustics, Middelfart, Denmark) and Telephonics TDH-39P earphones (Telephonics, NY, USA). Subjects were excluded from further testing if they did not meet the threshold criteria of $\leq$20 dB HL at octave frequencies from 250 – 8000Hz in both ears.

Psychophysical Testing

A double-walled, sound-attenuated booth was used for psychophysical testing. Auditory stimuli were presented diotically via Etymotic ER-2 headphones (Etymotic Research Inc., IL, USA). Following the design of Viemeister (1979), a 2-alternative forced choice (2AFC) task was used to determine psychophysical thresholds for AM detection. Each stimulus consisted of three continuous 500 ms periods of broadband noise (low-pass filtered at 8 kHz). In each trial, the first or third period was randomly chosen to be sinusoidally amplitude modulated. The second (centre) 500ms period was a reference period and always consisted of unmodulated broadband noise. Participants indicated which interval they thought contained modulation by means of a key press. Participants were tested separately at AM rates of 4 and 40Hz with 20 practice trials delivered at each rate before testing began. LEDs were used to indicate the two test intervals and feedback was given after each trial. The starting phase of the modulation was drawn randomly from a uniform distribution ranging from 0 to 2 π and the standard and comparison interval power were normalised to match the RMS of the centre period. Stimuli were generated at a sample rate of 44.1 kHz using Python and presented using PsychoPy (Peirce 2007) and the PyAudio module (`https://people.csail.mit.edu/hubert/pyaudio/`) through an EMU 0204 24-bit DAC (Creative Technology Ltd, Singapore). Stimuli were presented at

70 dB SPL as calibrated using an artificial ear (B&K 4153) and a sound level meter (B&K 2260). The AM depth of the comparison interval was systematically modified using an adaptive staircase procedure (Levitt 1971). Two interleaved staircases were used (1-up, 2-down; 1-up, 3-down). Before each trial, a random choice was made between the remaining staircases and this continued until each staircase had reached 14 reversals. A step size in modulation depth of 1 dB was used for the first 4 reversals of each staircase, which decreased to 0.5 dB for the next 4 reversals and to 0.25 dB for the final 6 reversals. Participants completed three runs (six staircases in total) at each of the 4 and 40Hz modulation rates. For 4 and 40Hz separately, an estimate of the 85% threshold from each run was calculated from the respective psychometric function (a cumulative gaussian) on a per-subject basis. The largest of the three modulation depth thresholds was discarded and the mean of the other two was used as the threshold depth for that subject in the TMS experiment.

fMRI Acquisition and Analysis

Subjects wore earplugs underneath the sound-attenuating headphones from the fMRI-compatible auditory stimulus delivery system (MR Confon, MR Confon GmBH). fMRI data were acquired using an Interleaved Silent Steady-State fMRI sequence (Schwarzbauer et al. 2006). The sound level of the scanner noise, not accounting for attenuation provided by earplugs and ear defenders, was 81 dB SPL during the quiet period and 98 dB SPL during the acquisition period. During the acquisition periods, whole head fMRI data (GE-EPI, TR = 2 s, TE = minimum full, FA = 90°) were collected using a GE Signa HDx 3T system (General Electric, WI, USA). A 128 × 128 pixel matrix with a field of view of 25.6 cm was used, giving an in-plane resolution of 2×2 mm. 24 interleaved slices were collected with a slice thickness of 2mm. Slices were oriented parallel with the Sylvian fissure while ensuring bilateral coverage of Heschl's gyrus and planum temporale.

Stimulus conditions were silence, noise, 4Hz AM noise, and 40Hz AM noise. All noise carriers were broadband (0 - 8000Hz) and the AM stimuli were fully modulated ($m = 1$). Stimuli were presented for 6 seconds followed by an acquisition period of 8s. Each stimulus was presented 6 times in each fMRI run and three runs were performed for each subject, each run lasting approximately 6 minutes. Data were analysed using Feat 5.98, part of FSL-4.1 (Smith et al. 2004), along with custom scripts which implemented filtering of the temporally non-contiguous data. Full details of this analysis procedure can be found in Hymers et al. (2015), with the exception that in this study spatial smoothing was performed using a kernel with a FWHM of 2mm.

Data from the multiple fMRI runs for each subject were combined using a fixed-effects analysis. A contrast of 4Hz and 40Hz AM noise combined over unmodulated noise was performed. The results were corrected for multiple comparisons using a cluster thresholding procedure ($Z > 2.3$, $p = 0.05$; Worsley 2001). TMS targets were defined individually for each subject using a superficial maxima voxel in a significant cluster on posterior superior temporal gyrus / planum temporale for each hemisphere.

Sagittal whole head structural T1-weighted data (3D FSPGR, TR = 7.8 ms, TE = minimum full, FA = 20°, Matrix 256 × 256, FOV = 29.0 cm, 176 slices, Slice thickness = 1 mm) were collected in the same scanning session as the fMRI data for participants who did not already have a structural MRI scan available.

TMS

A 2x2x3 repeated measures design was used during the main TMS experiment. Experimental factors were adaptation AM rate (2 levels; 4Hz, 40Hz), test AM rate (2 levels; 4Hz, 40Hz) and TMS (3 levels; sham, left, right). The dependent variable was accuracy on the AM detection task. Test AM rate was randomised within each session, so all subjects completed six TMS sessions, one for each adaptation AM rate by TMS combination. The order of sessions was approximately counterbalanced across subjects using a latin square design. Each TMS session was separated by at least 24 hours. A Brainsight frameless stereotaxic system (Rogue Research, Montreal, Canada) with chin rest and forehead support was used to monitor coil placement and ensure accurate TMS delivery. A Magstim Super Rapid^2 (Magstim, Whitland, UK) stimulator and a 80mm (external diameter of each ring) figure-of-eight coil (Magstim, Whitland, UK) was used for pulse delivery. The TMS coil was always oriented with the handle pointing horizontally in an anterior direction, parallel with the midline. Stimulation consisted of a double-pulse protocol with pulses separated by 50 ms. TMS pulse intensity was set to 65% of the maximum stimulator output (maximum = 2.6 T) based on previous state-dependent TMS studies (Cattaneo and Silvanto 2008; Cattaneo et al. 2008). This combination of TMS parameters is in line with published safety guidelines (Rossi et al. 2009). Before the first experimental TMS session, test pulses were delivered to ensure that subjects could tolerate this TMS protocol. The left and right TMS targets were obtained using the fMRI data as previously described. Sham TMS was performed by placing the coil at the vertex and rotating by 90°to remove any physiological effect of TMS but, crucially, retain peripheral factors such as the auditory artefact (Lisanby et al. 2001). Stimuli were delivered during the experiment using the same Etymotic ER-2 headphones and earpieces as were used during psychophysical testing.

A graphical outline of the TMS session can be seen in Figure 5.1. Every TMS session began with a period of auditory adaptation during which 10 minutes of broadband noise which had been fully amplitude modulated at a rate of either 4Hz or 40Hz was presented depending on the condition being tested. This period was to ensure participants were maximally adapted (Wojtczak and Viemeister 2005). All further behavioural adaptation within the session was at the same AM rate. A block of 150 trials was then performed. Each trial consisted of 4 seconds of behavioural adaptation followed by a 2AFC task. A double-pulse TMS burst was delivered during the adaptation sound, immediately prior to the 2AFC task. This timing is unusual for conventional TMS studies, but is consistent with previous state-dependent TMS studies. Each trial perceptually formed one continuous 5.5s auditory event. Figure 4.2 shows a single trial consisting of 4s of adaptation to 4 Hz noise followed by the 1.5s test period. The design of the 2AFC task was

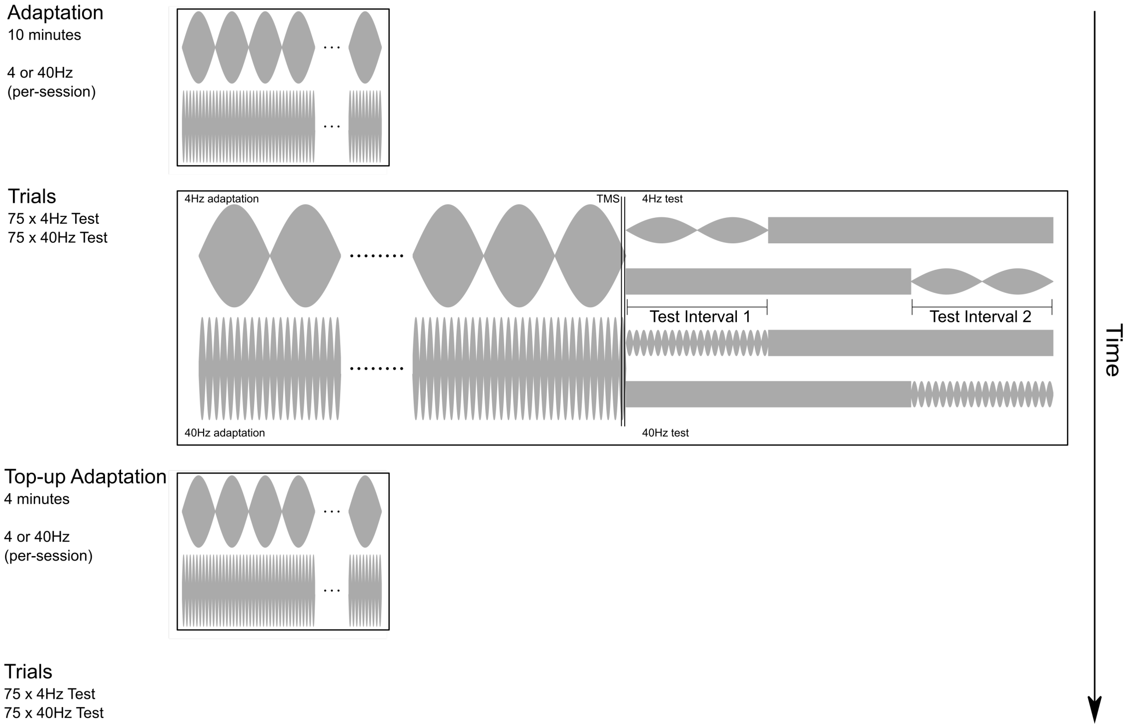

Figure 4.1: TMS session and trial layout. Time during the session runs from top to bottom. A given session used either 4Hz or 40Hz behavioural adaptation stimuli and one of three TMS targets (sham, left or right). Each session began with the subject listening to 10 minutes of fully modulated noise at the session adaptation rate. A block of 150 experimental trials was then performed; 75 test trials were modulated at 4Hz and the remaining 75 at 40Hz. Each trial consisted of 4s of adaptation noise (at the rate chosen for the session). Dual-pulse TMS stimulation was locked to the end of the trial adaptation stimulus. A test trial immediately followed the TMS pulses and consisted of a 2-AFC task - subjects were asked to select which of two intervals contained an AM noise. The AM depth was set based upon the initial psychophysical data for the subject. After 150 trials, subjects listened to four minutes of "top-up" adaptation noise. This was followed by a second block of 150 experimental trials of identical layout to the first block. The total number of TMS pulses in each session was 600.

identical to that used during baseline psychophysical testing, aside from the fixed (85% threshold) AM depth of the target interval. No feedback was given during the TMS sessions. During each block the test AM rate was pseudo-randomised; 75 trials were modulated at a rate of 4Hz and the other 75 were modulated at a rate of 40Hz. At the completion of the first block of trials, 4 minutes of fully modulated AM noise (top-up adaptation) were presented at the same rate used at the start of the experiment to ensure that the adaptation effect was consistent. This period also allowed the participant to have a short break from responding, during which the TMS coil was switched with an identical replacement to ensure the coil remained at an operational temperature. A second block of 150 trials was then completed.

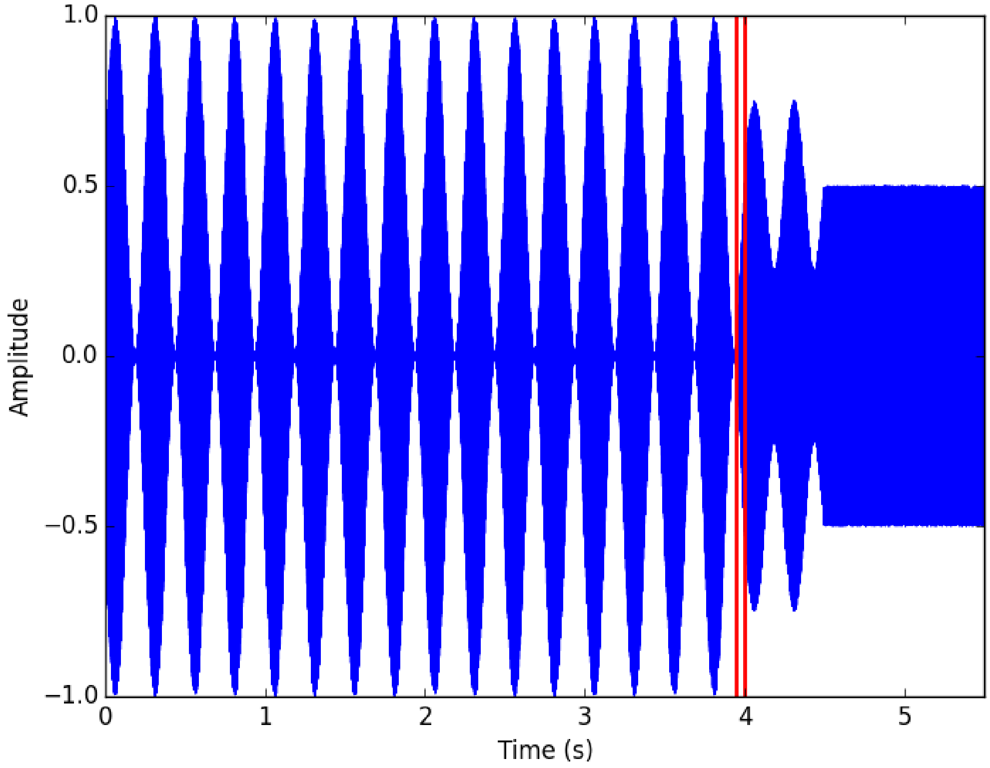

Figure 4.2: An example trial showing the 4s adaptation phase, directly followed by the 1.5s test phase. The adaptation and test rate in this example trial are both 4 Hz. The target period of the test phase is period 1. TMS pulse timing is indicated by the vertical red lines.

Key Methodological Changes

This section will detail the key similarities and differences between the design used in the study in Chapter 3 and the design used for this study. An AM detection task was used as a behavioural measure of AM sensitivity in the experiment reported in Chapter 3. While the overall task to detect the interval that was amplitude modulated remained the same in this study, the specific design was different. The key difference is that the 2-AFC trial in Chapter 3 consisted of five continuous 0.5 s noise periods, where either period 2 or 4 were modulated, however, the 2-AFC trial used here consisted of three continuous 0.5 s periods where either period 1 or 3 was modulated. Three periods were used in this task for consistency with the main TMS session, where the three period version of the task was used for efficiency.

The fMRI design used in this study was identical to the fMRI design reported in Chapter 3.

The key difference in TMS experimental design between the study reported in Chapter 3 and the study described here is the TMS pulse timing. TMS pulses were delivered during the 2-AFC trial in the experiment described in Chapter 3, due to the online TMS paradigm used. However, in the study reported here, the TMS pulses were delivered just prior to the 2-AFC test phase. This methodological change was required in order to recreate the conditions in which state-dependent TMS effects have been demonstrated.

The design of the auditory task within the TMS session also differed, primarily with the addition of the adaptation phase preceding each 2-AFC test phase. This adaptation period could consist of either 4 or 40 Hz AM noise and was included in order to reproduce the conditions required for a controlled state-dependent TMS adaptation effect. Each trial was longer in the experiment reported here compared to Chapter 3, due to this additional period of AM noise. Therefore, using a shorter overall 2-AFC test phase allowed more trials to be conducted in the same time, leading to increased efficiency.

Statistical Analysis

Generalised linear mixed-effects models (GLMM) were fit to the TMS accuracy data. This technique accounts for interdependence of the data arising from repeated measurements of the same participants; it also allows adjustments for non-normal distributions (Baayen et al. 2008; Dixon 2008; Jaeger 2008). A logit link function was used to account for the binomial distribution of the raw data. Separate models were fit to the 4Hz and 40Hz test AM rate data. Both models included fixed effects of TMS condition and behavioural adaptation condition and an interaction term. Participant was included as a random effect. Random intercepts and slopes for TMS condition and behavioural adaptation condition were estimated for each participant. The random effects structure was the maximal justified by the design (Barr et al. 2013); the model with maximal random effect structure was fit then systematically reduced until convergence. Treatment coding was used to measure the difference from the baseline condition of sham TMS with 4Hz AM adaptation (the baseline condition choice is arbitrary and was chosen to make other parameters easier to interpret). Finally, parametric bootstrapping was conducted on the model parameters to obtain 95% confidence intervals. All modelling was conducted using the *lme4* package (Bates et al. 2015) in *R* software (R Core Team 2017). Parametric bootstrapping was conducted using the *pbkrtest* package (Halekoh and Hojsgaard 2014).

4.4 Results

4.4.1 Psychophysics

The modulation depth at which 85% accuracy was achieved was calculated individually for each participant, for each AM test rate (4 and 40 Hz). The mean modulation depth across participants for the 4 Hz condition was -20.7 dB and for the 40 Hz condition was -21.2 dB. A paired t-test showed no significant difference between thresholds for 4 Hz and 40 Hz AM detection (t(7) = 0.888, p = 0.404).

4.4.2 fMRI derivation of TMS Targets

TMS targets in MNI-152 co-ordinate space for all participants can be found in Table 4.1.

Participant	Left hemisphere target	Right hemisphere target
1	-64, -22, 8	60, -20, 8
2	-52, -28, 6	62, -36, 6
3	-58, -18, 2	42, -18, 2
4	-54, -30, 2	64, -22, 6
5	-54, -30, 14	54, -22, 8
6	-40, -38, 14	66, -8, 2
7	-64, -28, 16	52, -34, 16
8	-38, -32, 14	44, -28, 12
Mean	-53.0, -28.3, 9.5	55.5, -23.5, 7.5

Table 4.1: MNI-152 co-ordinates for TMS targets in left and right hemispheres for all participants. The average co-ordinate across all participants is also given.

Figure 5.2 shows targets in both hemispheres for an example participant.

TMS target distance measurements were recorded for TMS pulses. This is a measure of the distance from the centre of the TMS coil to the fMRI-defined target. Four sessions lacked these measurements due to a technical issue. A linear mixed-effects model was fit to these data, with a fixed effect of TMS hemisphere and a random effect of participant. Parametric bootstrapping of the model parameters was conducted to obtain 95% confidence intervals. There was a negligible mean difference of -0.558 mm, 95% CI [-10.025, 8.631] between left hemisphere and right hemisphere conditions. There was no overall correlation between the AM task accuracy scores and the target distance measurements ($r = 0.008$, $p = 0.954$).

TMS positioning error measurements were also recorded. This is a measure of the difference between the fMRI-defined target and the estimated pulse trajectory. Four sessions lacked these measurements due to a technical issue. A linear mixed-effects model was fit to these data, with a fixed effect of TMS hemisphere and a random effect of participant. Parametric bootstrapping of the model parameters was conducted to obtain 95% confidence intervals. There was a negligible mean

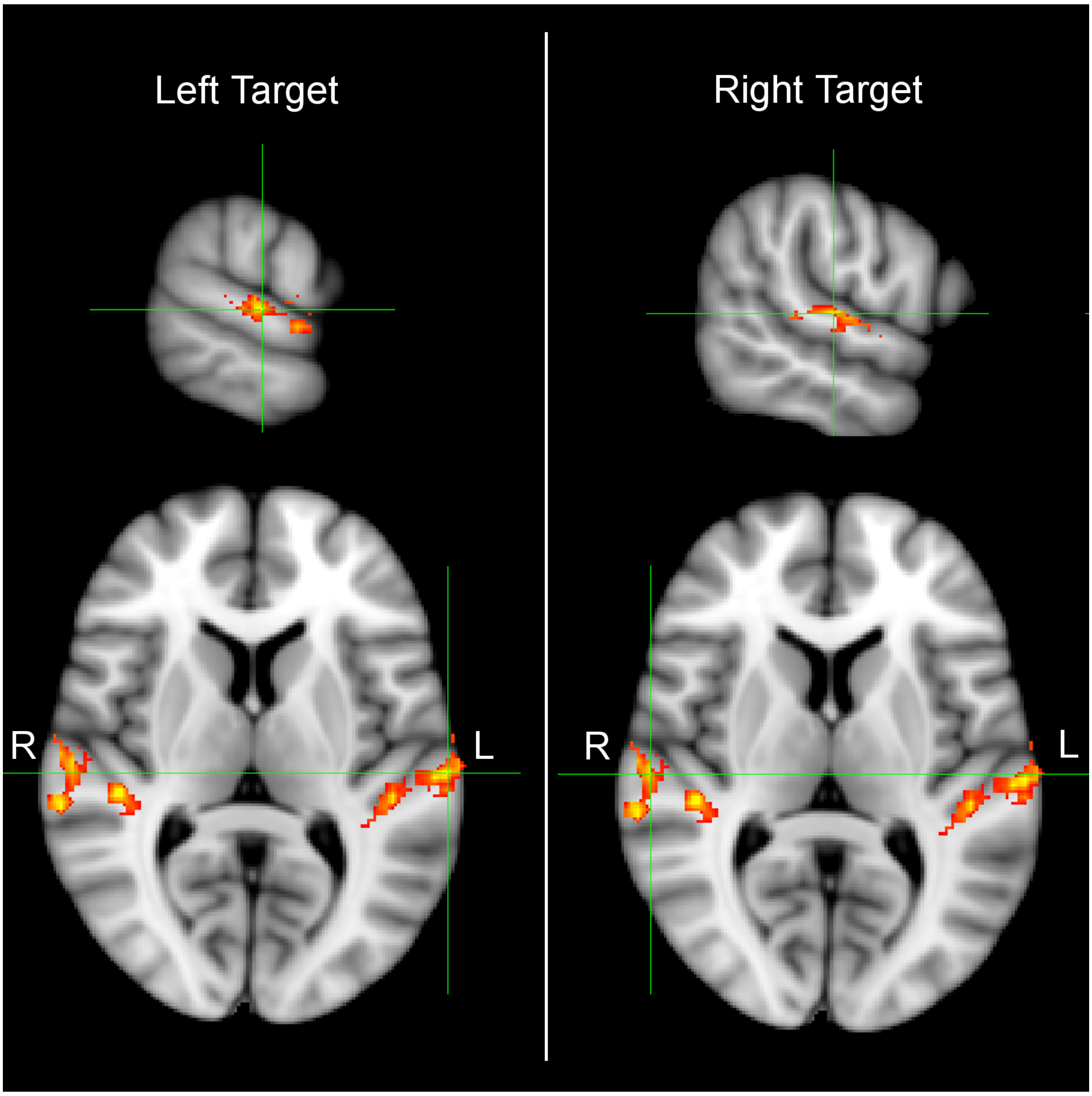

Figure 4.3: fMRI TMS target localisation data from an example participant. Overlay shows statistically significant clusters of activation (Z >2.3; p <0.05) in bilateral pSTG/PT in response to a contrast of 4 and 40 Hz sinusoidally amplitude modulated broadband noise conditions (combined) over unmodulated noise. TMS targets are indicated by the crosshairs. Functional data is shown here on the MNI-152 template brain - targets were transformed back into individual participant co-ordinate space. Target locations for this participant in MNI-152 co-ordinate space are: Left: $-64, -22, 8$ and Right: $60, -20, 8$

difference of -0.031mm, 95% CI [-0.082, 0.021] between left hemisphere and right hemisphere conditions. The overall mean error was 0.342mm, 95% CI [0.320, 0.364].

4.4.3 Sham TMS Behavioural Analysis

The effects of adaptation without TMS were assessed using the sham TMS conditions. This ensured that any noise from the TMS coil was incorporated into the baseline measure. For the 4 Hz test condition, a mean accuracy decrease of 26% was seen when participants were adapted to 4 Hz compared to when they were adapted to 40Hz. For the 40 Hz test condition, a mean accuracy decrease of 22% was seen when participants were adapted to 40 Hz compared to when they were adapted to 4 Hz AM noise. A paired t-test showed no significant difference in level of adaptation in the matched cases; i.e. 4 Hz adaptor followed by 4 Hz test versus 40 Hz adaptor followed by 40 Hz test ($t(7) = 1.222$, $p = 0.261$).

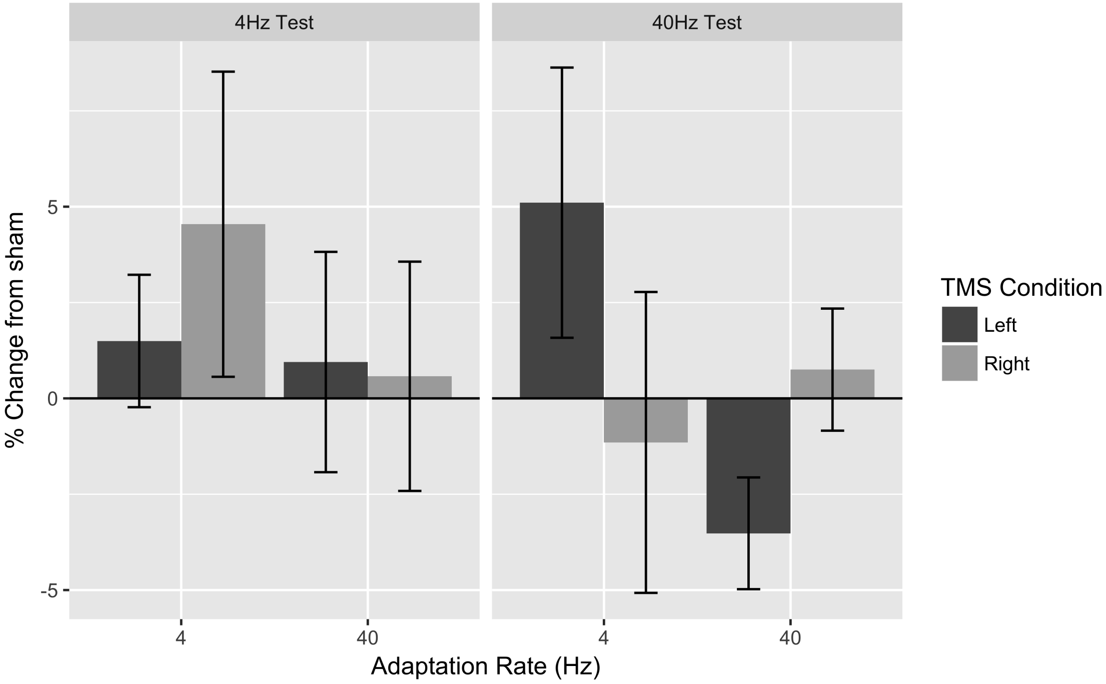

Figure 4.4: Difference in raw AM detection scores between sham TMS and left or right hemisphere TMS conditions are expressed here as percentage changes. Each bar represents the mean within-subject difference across all participants. Error bars represent the 95% confidence interval.

GLMM analyses

The data were divided based on test AM rate and separate generalised linear mixed effects models with logit link functions (GLMMs) were fit to the two subsets. These models assume that the log-odds of the binomial response variable varies linearly. Separate models were used for the 4 Hz and 40 Hz results due to convergence issues when a single model was used. Factors included in both final models were TMS condition (3 levels: sham, left hemisphere, right hemisphere) and adaptation condition (2 levels: 4 Hz and 40 Hz). Results are presented in logits to aid interpretation as confidence intervals are symmetrical about the point estimate and each coefficient value directly represents the effect size. Using logits also preserves the effect direction represented by the sign of the coefficients. The coefficient for the baseline condition represents the overall mean intercept and the other coefficients represent the (population-level) difference from this baseline for each condition. For consistency, the baseline condition for both models was sham TMS with adaptation to 4 Hz AM.

Whilst assessments of statistical significance using conventional metrics such as p-values are difficult to perform with GLMMs due to the unknown denominator degrees of freedom required to calculate the F-statistic, confidence intervals can be computed using parametric bootstrapping (Halekoh and Hojsgaard 2014). These 95% confidence intervals can be used to approximately assess statistical significance at the 0.05 alpha level. If a value of 0 is not contained within the upper and lower bounds of the confidence interval, then the result is approximately statistically significant using a type-I error rate of 5%. While there are clear issues surrounding the use and interpretation of dichotomous assessments of statistical significance (McShane et al. 2017), they are conventional and informative so the (approximately) statistically significant results have been indicated by an asterisk in the results tables.

GLMM analysis of 40 Hz test rate data

Results from this model are shown in Table 5.5 with approximately statistically significant results indicated by an asterisk. Model fit values are as follows: AIC = 376.7, BIC = 406.6 . Behavioural adaptation to 40 Hz AM had a clear negative effect on the 40 Hz test rate data, as shown by the negative *40Hz Adaptation* coefficient. This result replicates previous psychophysical work on selective adaptation to AM (e.g. Wojtczak and Viemeister 2005). TMS to the left hemisphere had a net positive effect on 40 Hz AM detection over both adaptation rates, as shown by the positive *Left Hemisphere TMS* coefficient. However, these main effects are confounded by a rate by hemisphere interaction, as seen in Figure 5.4. The negative *Left Hemisphere TMS x 40Hz Adaptation* interaction coefficient represents the differential effect that adaptation to 40 Hz AM had compared to 4 Hz adaptation, specifically in the left hemisphere TMS condition. As the baseline condition is 4 Hz AM with sham TMS, this interaction coefficient represents the net change from sham to left TMS and 4 Hz adaptation to 40 Hz adaptation.

GLMM analysis of 4 Hz test rate data

Results from this model are shown in Table 5.4 with approximately statistically significant results indicated by an asterisk. Model fit values are as follows: AIC = 389.1, BIC = 419.0 . Accuracy on the 4 Hz AM detection task was much lower when participants were behaviourally adapted to 4 Hz AM compared to when participants were behaviourally adapted to 40 Hz AM. This is shown by the positive *40Hz Adaptation* coefficient; the only coefficient from the 4 Hz test rate model where the 95% CI did not cross zero, as shown in Table 5.4. This is consistent with the aforementioned psychophysical work that showed peak adaptation for matched adaptation and test rates (Wojtczak and Viemeister 2005). The chosen baseline condition explains the difference in direction for the adaptation effect between the two models; a decrease in accuracy is shown in the 40 Hz GLMM whereas an increase in accuracy is shown in the 4 Hz GLMM. The lack of a statistically significant TMS or interaction effect suggests TMS was unable

	Dependent variable:			
	Accuracy	95% CI		
Baseline	1.018	0.669	1.378	
Left Hemisphere TMS	0.302	0.096	0.519	$*$
Right Hemisphere TMS	-0.068	-0.353	0.197	
40Hz Adaptation	-1.014	-1.371	-0.663	$*$
Left Hemisphere TMS x 40Hz Adaptation	-0.446	-0.712	-0.202	$*$
Right Hemisphere TMS x 40Hz Adaptation	0.099	-0.154	0.356	

Table 4.2: GLMM results for the 40 Hz AM test conditions. Accuracy values and 95% CIs are in logit units. 'Baseline' represents the mean in the baseline condition of sham TMS with 4 Hz behavioural adaptation. All other coefficients represent the difference from this baseline, i.e. the effect size. For example, the 40Hz Adaptation co-efficient shows the effect of 40Hz behavioural adaptation compared with the baseline of 4Hz behavioural adaptation; in this case (40Hz test trials), the co-efficient shows the expected negative direction, indicating a decrease in AM detection accuracy when participants were behaviourally adapted to 40Hz. An asterisk is used to indicate approximately statistically significant results.

to affect AM processing in the same way as with 40 Hz processing.

	Dependent variable:			
	Accuracy	95% CI		
Baseline	0.438	0.182	0.719	
Left Hemisphere TMS	0.068	−0.167	0.298	
Right Hemisphere TMS	0.205	−0.036	0.422	
40Hz Adaptation	1.437	1.195	1.679	∗
Left Hemisphere TMS x 40Hz Adaptation	0.023	−0.293	0.310	
Right Hemisphere TMS x 40Hz Adaptation	−0.137	−0.434	0.147	

Table 4.3: GLMM results for the 4 Hz AM test conditions. Accuracy values and 95% CIs are in logit units. 'Baseline' represents the mean in the baseline condition of sham TMS with 4 Hz behavioural adaptation. All other coefficients represent the difference from this baseline, i.e. the effect size. For example, the 40Hz Adaptation co-efficient shows the effect of 40Hz behavioural adaptation compared with the baseline of 4Hz behavioural adaptation; in this case (4Hz test trials), the co-efficient shows the expected positive direction, indicating an increase in AM detection accuracy when participants were behaviourally adapted to 40Hz. An asterisk is used to indicate approximately statistically significant results.

4.5 Discussion

The present study was designed to investigate whether TMS to auditory areas could affect AM detection performance and if so, whether the induced behavioural changes showed state-dependency and/or provided support for asymmetric accounts of auditory processing. Three specific research questions were posed.

The first question asked whether AM detection could be affected by TMS to auditory cortex. A significant effect of left hemisphere TMS was shown compared to sham TMS, but only in the 40Hz AM detection condition.

The second question asked whether TMS effects showed state-dependency, i.e. did behavioural adaptation have an effect on the direction or degree to which

TMS modulated performance on the task. In the conditions with left hemisphere TMS and 40Hz AM detection, state-dependent TMS effects were found, though the change in behavioural performance was not in the predicted direction; a facilitation of AM detection was predicted in matched adaptation conditions, however in the matched 40 Hz condition, TMS caused a decrease in AM detection accuracy.

The final question asked whether any TMS effects were consistent with asymmetric models of auditory processing, with directional predictions following the AST hypothesis (Poeppel 2003). Left hemisphere stimulation was predicted to preferentially modulate 40Hz AM detection whilst right hemisphere stimulation was predicted to preferentially modulate 4Hz AM detection. Left hemisphere stimulation was shown to affect the detection of 40Hz AM, whilst no effect of right hemisphere stimulation was shown on the detection of either 4 or 40Hz AM.

4.5.1 Effect of AM rate

Our results show that whilst TMS to the left hemisphere affected detection of 40Hz AM, neither right nor left hemisphere TMS affected detection of 4Hz AM. Before considering possible explanations for this difference, some potential confounds must be addressed. The number of modulation cycles within the presentation period is known to affect AM detection ability (Sheft and Yost 1990). The 500ms period contained 2 cycles of AM in the 4Hz test condition compared to 20 cycles in the 40Hz test condition. To address this possible confound, detection difficulty was calibrated separately for each AM rate and for each participant.

Several studies have shown that BOLD signal changes are both smaller and less distributed in response to higher frequency AM compared to lower frequency AM (Giraud et al. 2000; Harms and Melcher 2002). Therefore, the choice of TMS target locations could also affect whether, and to what degree, processing of different modulation rates would be affected by TMS. Cortical TMS targets were chosen based on fMRI responses to 4Hz and 40Hz AM combined, contrasted over unmodulated noise to avoid biasing the targets towards areas more specialised for processing either rate.

The observation that TMS only affected 40Hz AM detection could be due to the use of a different and possibly more robust mechanism for coding lower rates of modulation (such as 4Hz). Evidence from multiple modalities has converged in support of such an account where low and high frequency AM is coded by different mechanisms. fMRI studies have found that lower frequency AM elicits a more sustained response whilst higher frequency AM produces more phasic responses at onset and offset of the stimulus, possibly representing a single auditory event (Giraud et al. 2000; Harms et al. 2005). Direct cortical recordings using invasive intracranial electrodes have also provided strong support for a shift in coding mechanism as AM rate increases (Brugge et al. 2009; Nourski et al. 2012; Nourski 2017). Non-invasive electrophysiological evidence suggests that the balance of phase-locked (PL) and non phase-locked (NPL) activity changes as AM rate increases; PL activity dominates at lower rates whilst more NPL activity is observed with increasing AM rate (Tang et al. 2016; Wang 2018). The cutoff point for this transition from mostly PL to mostly NPL activity is estimated to be 50Hz (above the rates studied here), but the amount of NPL activity does differ between 4Hz and 40Hz. Though PL activity still dominates the response to 40Hz AM, Tang et al. (2016) found a significant increase in NPL activity over that measured for 4Hz AM. It is possible that this difference in coding mechanisms might contribute to the observed results. Any mechanism which relies on coding the onset and offset of a stimulus could be more susceptible to TMS than one reliant on sustained responses. If, for example, the single cortical event coding for the onset of 40Hz AM was disrupted, this may have a significant impact, particularly at short stimulus durations in which there is little time for the system to recover. In comparison, a sustained system where PL neurons fire in synchronisation with slower modulations would likely be more robust to a single disruptive event. Further, it would be biologically useful for the system that processes low-frequency AM it to be more robust to disruption due to its importance for speech intelligibility. This explanation is challenged however, by the classic finding that with sustained periods of auditory stimulation, peak PL responses are shown for 40Hz AM; the 40Hz auditory steady-state response (Galambos et al. 1981; Ross et al. 2000). Though PL and NPL activity is evident during 40Hz processing, the role and relative importance of each mechanism is not clear. The suggestion that TMS may differentially affect disparate neuronal coding mechanisms has important implications for the design of future TMS experiments and warrants further investigation.

4.5.2 Hemispheric differences

Before considering explanations of the changes in behavioural performance caused by TMS to the left and right hemisphere, we can first rule out several methodological confounds. Neither the distance to the TMS target, nor the accuracy with which TMS was delivered differed significantly between the two hemispheres or between test conditions. In addition, there was no significant correlation across subjects between task accuracy scores and the distance to target, suggesting that any distance effects can be discounted.

Previous research using transcranial direct current stimulation (TDCS) to investigate auditory temporal resolution using a gap detection task has shown differential susceptibility of the left and right hemispheres to stimulation. A detrimental effect on performance was found exclusively when anodal TDCS was applied to the left hemisphere (Heimrath et al. 2014). It is interesting that the current results accord with this finding as there is evidence that gap detection tasks and estimation of the temporal modulation transfer function (TMTF; using the method employed in this study, though the full TMTF was not estimated here) are measuring different aspects of auditory temporal acuity (Shen 2014). Disruptive TMS effects on duration discrimination have been shown exclusively for the right hemisphere (Bueti et al. 2008). However, these studies differ not only in both the nature of the behavioural task, but also the nature of the stimulus. Whilst Bueti et al. (2008) used tonal stimuli, Heimrath et al. (2014) and the current study use relatively broadband noise stimuli. A question of interest is therefore whether the spectral nature of the stimulus has an influence on hemispheric lateralisation of function and how this interacts with the specific behavioural task under investigation.

4.5.3 State-dependent effects

The initial prediction for this study was that TMS combined with behavioural adaptation at a modulation rate matched to the test modulation rate would lead to an improvement in AM detection accuracy. Instead, an unexpected dissociation was shown between performance on the 40Hz AM detection task in the matched and un-matched adaptation conditions: a behavioural facilitation was observed in the unmatched condition (4 Hz adaptation) whilst a reduction in performance was seen in the matched condition (40 Hz adaptation). An overarching framework that accounts for both online and state-dependent effects has recently been proposed (Silvanto and Cattaneo 2017). It is possible to reconcile these results within this framework, however not using a purely state-dependent TMS explanation. This suggests that modifications made to the experimental design when converting it to investigate auditory function may have affected the nature of the disruption. The majority of state-dependent TMS experiments have used visual tasks involving brief *simultaneous* presentations of multiple stimuli with TMS pulses delivered on stimulus onset (Silvanto et al. 2007; Cattaneo and Silvanto 2008; Silvanto and Muggleton 2008b). To account for the necessarily more temporally extended nature of auditory stimuli, modifications were required when designing this auditory state-dependent TMS paradigm. In each trial, an auditory adaptation period immediately preceded a 2-AFC decision, with TMS delivered just prior to the onset of the 2-AFC task. In matched adaptation trials, though the stimulation is delivered whilst the participant is adapted, it is actually delivered *during* AM processing at the target rate, therefore corresponding more closely to online TMS. The suprathreshold TMS would then be expected to impair performance as the neurons tuned to the target rate have been shifted into their inhibitory range (Silvanto and Cattaneo 2017: fig.1). This key timing difference could explain the ‘virtual lesion’ style detrimental TMS effects seen in the matched 40Hz condition. In unmatched adaptation trials, participants received TMS during processing of one rate of AM and were subsequently tested at the other rate.

Within this same overarching framework, two possible explanations exist that lead to a facilitation of behavioural performance. At the point at which TMS was delivered, neural populations responsible for processing the test rate would be suppressed relative to populations responsible for processing the adaptation rate. This would lead to test rate populations being in their facilitatory range when activated by the suprathreshold TMS, leading to behavioural improvements at detecting 40Hz AM (Silvanto and Cattaneo 2017: fig. 2). Alternatively, It is also possible that the neural populations that process 40Hz AM may have been adapted by the 4Hz AM. Maximal adaptation, as measured by decrease in detection ability, occurs when adaptor and test rates are identical, however there is evidence for a graded decrease in adaptation as the difference between the two rates increases (Houtgast 1989; Bacon and Grantham 1989; Wojtczak and Viemeister 2005). Based on the estimated shape of the modulation filter with centre frequency of 40Hz, it may be expected that populations involved in processing 40Hz would be adapted by exposure to 4Hz AM. Further, the estimated shape of modulation filters with centre frequencies of 4 and 40Hz differs such that the 4Hz filter is much more narrow (5Hz bandwidth), suggesting an asymmetry where populations optimised for coding 4Hz would not be adapted by exposure to 40Hz AM, but populations optimised for coding 40Hz may still be adapted somewhat by exposure to 4Hz AM (Dau et al. 1997a). Under this assumption, the outcome in the unmatched condition would follow state-dependent TMS predictions, where adaptation has suppressed the activity of 40Hz populations, shifting the inhibitory/facilitatory range of these populations such that suprathreshold TMS leads to behavioural facilitation. With the current dataset, it is not possible to dissociate these two mechanisms, both of which would lead to an increase in performance for 40Hz AM detection.

Overall, the timing of TMS delivery is central to this explanation of the observed results. The importance of this parameter has been previously noted (Silvanto and Pascual-Leone 2008), with differential effects shown when pulses were delivered before and during task performance across different studies, with and without adaptation. In online TMS studies, no adaptation is used and supra-threshold TMS is delivered during the test stimulus, causing a decrease in performance. The current study therefore demonstrates that TMS can cause either disruptive or facilitatory effects on 40Hz AM detection depending on which AM rate is perceived during TMS delivery. Further investigating how TMS at different timepoints affects AM detection will help to elucidate the mechanisms of this effect and improve the utility of auditory TMS as an investigative tool.

4.6 Conclusions

This study showed, for the first time, that low-level mechanisms involved in AM detection, a key component of speech processing, can be affected using TMS. This important finding demonstrates that TMS can be a useful technique for investigating low-level auditory processing mechanisms and opens the door to new research in this area. It was also shown that the TMS effect on AM detection shows state-dependency, i.e., the underlying state of the AM detection system when TMS is delivered has a differential effect on the outcome. It was further demonstrated that systematically modifying this underlying state using behavioural adaptation is an effective way to increase the functional resolution of TMS and alter the nature of TMS effects on AM detection. The results of this study provide support for the importance of the left posterior superior temporal gyrus for processing 40Hz AM. It also suggests that 4Hz AM may be encoded using a different mechanism that is more robust to interference from TMS. This illustrates a potential limitation of the method and would benefit from further investigation. Finally, in addition to the direct implications for auditory processing, methodological considerations for the use and interpretation of state-dependent TMS were discussed so that important factors such as subtle timing and task manipulations are accounted for in future work. This study demonstrated the utility of TMS as a new and potentially valuable tool for furthering understanding of low-level auditory processing in health and disease.

4.7 Acknowledgments

AMP is supported by a PhD studentship from the Economic and Social Research Council (grant number: ES/J500215/1). The authors would like to acknowledge assistance from Faith Marsh, Elenor Morgenroth and Sofia Tsitsopoulou with the collection of the data as part of their MSc empirical research project.

Chapter 5

Amplitude Modulation Depth Discrimination is not Affected by Continuous Theta Burst Stimulation to Left or Right Auditory Cortex

5.1 Abstract

Low frequency amplitude modulations (AM) that comprise the speech envelope are a critical component for the intelligibility of a speech signal. The differential contribution of left and right auditory areas to the processing of these low-level speech-relevant auditory stimuli remains unclear. Multiple plausible models have been suggested to account for observed hemispheric asymmetries in processing many auditory stimuli. AM depth discrimination is a crucial aspect of auditory processing as it determines the smallest level of information that can be decoded from an amplitude modulated signal such as speech (Wakefield and Viemeister 1990). It is unknown whether sensitivity to AM depth is susceptible to disruption by non-invasive brain stimulation, however, this could be a novel and effective tool to further investigate AM processing. In order to investigate this, a functional magnetic resonance imaging (fMRI)-guided transcranial magnetic stimulation study was conducted. AM depth discrimination thresholds for sinusoidally amplitude modulated broadband noise (0 – 8 kHz) stimuli were obtained at two AM rates; 4 and 40 Hz. Cortical locations involved in the processing of AM were then isolated for each participant using an interleaved silent steady state fMRI paradigm. An AM depth discrimination task was then completed at threshold pre- and post- continuous theta-burst stimulation (cTBS). Left auditory cortex, right auditory cortex and sham TMS conditions were included to investigate any asymmetrical effects, as may be predicted by asymmetric models and previous TMS work (Poeppel 2003; Andoh et al. 2015). Multilevel modelling results showed no evidence of statistically significant differences for any experimental comparisons. Possible explanations for this are considered and recommendations for future investigations are made.

5.2 Introduction

The auditory system is able to continuously track ongoing modulations of auditory signals in remarkably high resolution; an ability that is particularly important for decoding speech. Speech consists of highly complex concurrent modulations of both amplitude and frequency, yet the human auditory system decodes it with relative ease. The exact mechanisms underlying speech comprehension are still unknown, however, the relative contributions of disparate low-level aspects have been investigated and amplitude modulations (AM) of the speech envelope have been shown to be particularly important (Greenberg 2004; Joris et al. 2004). In fact, low frequency AM (<16 Hz) is known to be critical for speech to be intelligible (Rosen 1992; Drullman et al. 1994a,b) and even with severely degraded spectral information, adequately preserved AM can be sufficient for intelligibility (Shannon et al. 1995). The ability to detect these modulations varies between individuals and poor temporal processing is common in disorders that include a language processing deficit, such as dyslexia (Lehongre et al. 2013; Cutini et al. 2016). Understanding the neural mechanisms involved in this process is a key stage in fully characterising these disorders and developing new treatments. With this aim, measures of temporal resolution such as gap detection and the temporal modulation transfer function have been developed (Fitzgibbons and Wightman 1982; Viemeister 1979). While these measures have clinical utility, it is not clear exactly what aspects of temporal processing they are each tapping into (Kumar et al. 2012; Shen 2014).

AM depth discrimination is the ability to detect a difference in the AM depth of two amplitude modulated stimuli. This is a particularly interesting measure as it reflects the smallest degree of AM required to decode a unit of information (Wakefield and Viemeister 1990). AM in natural speech is usually suprathreshold, so this measure more closely replicates the processes involved in speech decoding than basic AM detection. (Schlittenlacher and Moore 2016) While both AM detection and AM depth discrimination are measures of auditory temporal sensitivity, they are dissociable, for example, it was demonstrated that hearing impaired participants outperform normal-hearing participants on AM detection (at 30dB SPL) but perform more poorly than normal-hearing participants at AM depth discrimination (Schlittenlacher and Moore 2016). This suggests that the neural mechanisms underlying AM detection and AM depth discrimination are distinct as they respond differently to peripheral insult. The exact mechanisms underlying these functions are largely unknown, however the cortical system, at least in part, has been localised to secondary auditory cortical regions. Functional magnetic resonance imaging (fMRI) activations to AM are commonly measured in posterior superior temporal gyrus (pSTG) / planum temporale (PT), suggesting that these areas are part of the system that processes AM (Giraud et al. 2000; Harms and Melcher 2002; Griffiths and Warren 2002). Despite this consistent finding, the macrostructure of auditory areas is generally highly variable inter-individually and asymmetrical between hemispheres intra-individually (Geschwind and Levitsky 1968; Abdul-Kareem and Sluming 2008; Liem et al. 2014: e.g.). In complement to the observed macrostructural asymmetries in auditory cortical regions, behavioural and neuroimaging evidence is also mounting for functional asymmetries (Giraud et al. 2000; Boemio et al. 2005; Luo and Poeppel 2012; Abrams et al. 2008; Morillon et al. 2010; Han and Dimitrijevic 2015). Further, these asymmetries have behavioural relevance, for example, abnormal patterns of asymmetry have been linked to poor reading skills in individuals without language disorders (Abrams et al. 2009), however the true underlying nature of these asymmetries is contentious (Poeppel 2014).

One explanation for the hemispheric asymmetries holds that left hemisphere auditory areas are specialised for processing rapid temporal changes in the auditory signal whereas the right hemisphere is more specialised for processing spectral content (Zatorre and Belin 2001; Zatorre et al. 2002). Informed by evidence for distinct frequencies of cortical oscillatory activity during speech processing, a co-existing account focusses on the different timescales of processing between the two hemispheres. The Asymmetric Sampling in Time hypothesis holds that the left and right auditory cortices integrate the incoming signal using sampling windows of different lengths (Poeppel 2003; Giraud and Poeppel 2012). This results in a hemispheric asymmetry where left auditory cortex is specialised for processing high frequencies and right auditory cortex is specialised for processing low frequencies. The lengths of these sampling windows (~ 20 and ~ 250 ms, respectively) align with key frequencies in the speech signal; the syllabic (~ 4 – 8 Hz) and phonemic rate (~ 40 Hz). This model would predict that AM depth discrimination at an AM rate of 40 Hz would be preferentially processed by the left auditory cortex and AM depth discrimination at an AM rate of 4 Hz would be preferentially processed by the right auditory cortex. However, while dichotomous distinctions like these are appealing, they are often reductionist and may overlook key aspects, such as the true nature of the speech signal. It has been suggested that the left hemisphere may show a specialisation for speech processing while the right may not (McGettigan and Scott 2012).

Non-invasive brain stimulation methods are a relatively novel group of techniques that allow researchers to directly manipulate ongoing brain activity and can provide causal evidence to link cortical structures with many functions (Wagner et al. 2009). Transcranial magnetic stimulation (TMS) involves inducing a time-varying magnetic field to temporarily and non-invasively modulate activity in targeted neural systems (Barker et al. 1985). TMS has been successfully applied to investigate the neural mechanisms underlying many cognitive functions relating to speech and language (Hartwigsen 2015), and has also proved useful for investigating lower level auditory functions such as melody discrimination (Andoh and Zatorre 2011). Further, while online TMS to auditory areas has been shown to affect AM detection ability, the feasibility of TMS to perturb AM depth discrimination has not been investigated.

Continuous theta-burst stimulation (cTBS) is an offline, repetitive TMS protocol where clustered bursts of low intensity pulses (50 Hz bursts of 3 pulses, separated by 200 ms) are continuously delivered in a train lasting up to 40 seconds. The low intensity of stimulation allows for the safe delivery of many more pulses than standard repetitive TMS protocols in a shorter time frame. cTBS was designed to mimic the mechanism underlying neural long-term depression has been shown to produce suppressive aftereffects lasting up to one hour (Huang et al. 2005). The high efficiency and tolerability of the method has led to its application with clinical populations including those with disorders with auditory components, such as auditory hallucinations (Ray et al. 2015) and tinnitus (Weisz et al. 2012). Although cTBS is most commonly applied to modulate motor cortex excitability (Chung et al. 2016; Suppa et al. 2016), it is also commonly used as a basic science tool to investigate cognitive processes using behavioural measures. In the auditory behavioural domain, however, there is a paucity of previous applications of the method. Most notably, cTBS has been applied in combination with fMRI to further investigate functional asymmetries and the temporary changes cTBS causes to the underlying functional connectivity profile (Andoh and Zatorre 2013; Andoh et al. 2015). These studies found that stimulating right antero-lateral Heschl's gyrus led to decreased performance in a melody discrimination task, however this was linked to increased contralateral homologous activity, suggesting interhemispheric compensatory effects may be involved (Andoh and Zatorre 2013). It was also shown that right hemisphere cTBS causes more widespread decreases in functional connectivity compared with cTBS to the left hemisphere (Andoh et al. 2015). Although cTBS is typically considered a disruptive paradigm (due to suppressed EMG responses when applied to motor areas), these studies show that auditory networks respond very differently and cTBS can lead to facilitatory effects on auditory behavioural tasks. This type of conflicting result is common in TMS studies and the direction of effects can be difficult to predict (Weisz et al. 2012; López-Alonso et al. 2014). For example, one study investigating the efficacy of cTBS for treating tinnitus found significant effects of cTBS for all participants, however, the direction of the effect differed (Müller et al. 2013). TMS effects can also vary in magnitude or direction based on the underlying state of the cortex when TMS is delivered (Silvanto and Muggleton 2008a; Cattaneo et al. 2008; Silvanto et al. 2018). Despite this variability, cTBS is particularly well-suited for studying auditory perception. Primarily, this is because stimuli can be presented after the stimulation, during the aftereffect period. In the aftereffect period, there are no auditory TMS artefacts that could potentially mask the stimuli. Further, the efficiency of cTBS and the potential to induce longer lasting changes may be beneficial for future low-level auditory investigations. However, it is not currently known whether cTBS is an effective methods for modulating AM processing. These three principal reasons provided the rationale for using cTBS to investigate AM processing. Though cTBS has been successfully applied to probe motor function and has been shown to have effects on auditory neural networks, it has not been used to investigate processing of AM noise.

The current study combined auditory psychophysics and fMRI-guided TMS to further investigate the neural systems underlying AM depth discrimination. AM depth discrimination thresholds were measured for sinusoidal AM broadband noise at rates of 4 and 40 Hz. TMS targets were obtained for each participant using peak BOLD fMRI activity to AM at rates of 4 and 40 Hz. AM depth discrimination ability was then tested following cTBS to left or right auditory cortex and compared with a pre-TMS baseline. A sham TMS condition was also included to control for the influence of peripheral effects such as the auditory TMS click artefact, without any magnetic stimulation effects. It was, however, unclear if the complex click train produced during cTBS might affect perception of subsequent AM and the inclusion of this condition allowed the efficacy of this sham method to be tested.

The following three experimental hypotheses were proposed.

- Firstly, it was predicted that cTBS to both hemispheres would affect AM depth discrimination ability when compared to a pre-TMS baseline.
- Secondly, it was predicted that cTBS to the right hemisphere would have a largest disruptive effect on 4 Hz AM depth discrimination.
- Finally, it was predicted that cTBS to the left hemisphere would have a largest disruptive effect on 40 Hz AM depth discrimination.

5.3 Materials and Methods

5.3.1 Subjects

Ten subjects (5 female; mean age = 22.5 years, SD = 4.65) participated in the experiment. Subjects responded to an advert and were paid or reimbursed in course credit for participation except for one participant who was a member of the research team. All subjects were TMS and MRI eligible, right-handed, native English speaking and had self-reported normal hearing. The project was approved by the Research Governance Committee, York Neuroimaging Centre, University of York and conformed to the guidelines of the Declaration of Helsinki. All participants gave written informed consent.

5.3.2 Experimental Design

Audiogram

Subjects underwent an air conductance pure tone audiogram to determine that hearing thresholds were above eligibility requirements. Audiograms were conducted using an Interacoustics AD226 audiometer (Interacoustics, Middelfart, Denmark) and Telephonics TDH-39P earphones (Telephonics, NY, USA). Subjects were excluded from further testing if they did not meet the threshold criteria of $\leq$20 dB HL at octave frequencies from 250 – 8000Hz in both ears.

Psychophysical Testing

A double-walled, sound-attenuated booth was used for psychophysical testing. Auditory stimuli were presented diotically via Sennheiser HD650 headphones (Sennheiser, Wedemark, Germany). The two-alternative forced choice (2AFC) task used by Wakefield and Viemeister (1990) was used to determine psychophysical thresholds for AM depth discrimination. Each stimulus consisted of three continuous 500 ms periods of broadband noise (low-pass filtered at 8 kHz). In each trial, both the first and third period were sinusoidally amplitude modulated. The second (centre) 500ms period was always unmodulated. This stimulus forms one continuous 1.5 s percept. Either the first or third period was always modulated at a standard depth of 0.36m but the AM depth of the other period (comparison period) was varied. Participants indicated which interval (either interval one or three) they thought was modulated at a greater depth by pressing the left or right arrow key on a standard keyboard. Participants were tested separately at AM rates of 4 and 40Hz with 20 practice trials delivered at each rate before testing began. The starting phase of the modulation was chosen randomly from a uniform distribution ranging from 0 to 2 π and the standard and comparison interval power were normalised to match the RMS of the centre period.

The AM depth of the comparison interval was systematically modified between

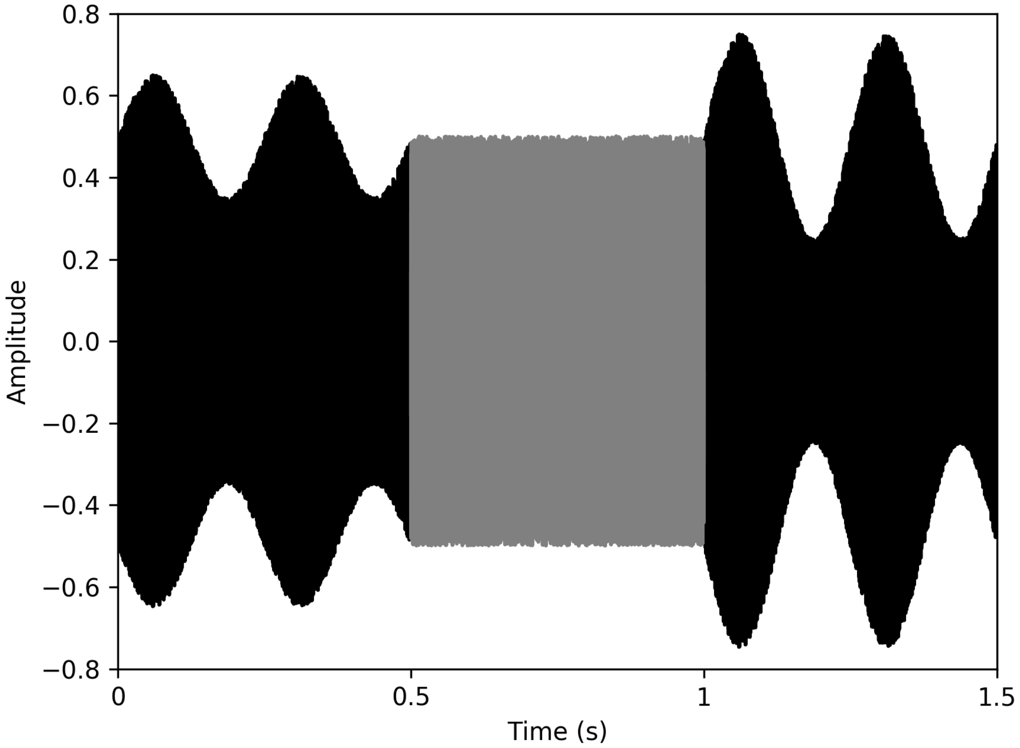

Figure 5.1: This example trial could be from either the psychophysical testing or main task run. In the psychophysical testing sessions, the difference in AM depth between period one and period three varied between trials. During the main task runs, this depth difference remained constant (set per participant at their 70% threshold value) but direction from the standard depth differed. The AM rate (4 or 40 Hz) and TMS target (left, right or sham) did not differ within a session.

trials using an adaptive staircase procedure (Levitt 1971). Four interleaved staircases were used (1-up, 2-down; 1-up, 3-down; 2-up, 1-down, 3-up, 1-down). This method allowed measures of sensitivity for AM depths both greater and less than the standard depth ($0.36m$). Before each trial, a random choice was made between the remaining staircases and this continued until each staircase had reached 14 reversals. A step size of 1 dB was used for the first 4 reversals of each staircase, which decreased to 0.5 dB for the next 4 reversals and to 0.25 dB for the final 6 reversals. Participants completed three runs at each of the 4 and 40Hz modulation rates. For 4 and 40Hz separately, the modulation depth delta where 70% accuracy was achieved was estimated from each run from the respective psychometric function on a per-subject basis. The largest of the three modulation depth delta estimates was discarded and the mean of the remaining two was used as the threshold depth for that subject in the TMS sessions. All stimuli were generated at a sample rate of 44.1 kHz using Python and presented using PsychoPy (Peirce 2007) and the PyAudio module (`https://people.csail.mit.edu/hubert/pyaudio/`) through an EMU 0204 24-bit DAC (Creative Technology Ltd, Singapore). Stimuli were presented at 70 dB SPL as calibrated using an artificial ear (B&K 4153) and a sound level meter (B&K 2260).

fMRI Acquisition and Analysis

Subjects wore earplugs underneath the sound-attenuating headphones from the fMRI-compatible auditory stimulus delivery system (MR Confon, MR Confon GmBH). fMRI data were acquired using an Interleaved Silent Steady-State fMRI sequence (Schwarzbauer et al. 2006). The sound level of the scanner noise, not accounting for attenuation provided by earplugs and ear defenders, was 81 dB SPL during the quiet period and 98 dB SPL during the acquisition period. During the acquisition periods, whole head fMRI data (GE-EPI, TR = 2 s, TE = minimum full, FA = 90°) were collected using a GE Signa HDx 3T system (General Electric, WI, USA). A 128 × 128 pixel matrix with a field of view of 25.6 cm was used, giving an in-plane resolution of 2 × 2 mm. 24 interleaved slices were collected with a slice thickness of 2 mm. Slices were oriented parallel with the Sylvian fissure while ensuring bilateral coverage of Heschl's gyrus and planum temporale.

Stimulus conditions were silence, noise, 4Hz AM noise, and 40Hz AM noise. All noise carriers were broadband (0 – 8000Hz) and the AM stimuli were fully modulated ($m = 1$). Stimuli were presented for 6 seconds followed by an acquisition period of 8s. Each stimulus was presented 6 times in each fMRI run and three runs were performed for each subject, each run lasting approximately 6 minutes. Data were analysed using Feat 5.98, part of FSL-4.1 (Smith et al. 2004), along with custom scripts which implemented filtering of the temporally non-contiguous data. Full details of this analysis procedure can be found in Hymers et al. (2015), with the exception that in this study spatial smoothing was performed using a kernel with a FWHM of 2mm.

Data from the multiple fMRI runs for each subject were combined using a fixed-effects analysis. A contrast of 4Hz and 40Hz AM noise combined over unmodulated noise was performed. The results were corrected for multiple comparisons using a cluster thresholding procedure ($Z > 2.3$, $p = 0.05$; Worsley 2001). TMS targets were defined individually for each subject using a superficial maxima voxel in a significant cluster on posterior superior temporal gyrus / planum temporale for each hemisphere.

Sagittal whole head structural T1-weighted data (3D FSPGR, TR = 7.8 ms, TE = minimum full, FA = 20°, Matrix 256 × 256, FOV = 29.0 cm, 176 slices, Slice thickness = 1 mm) were collected in the same scanning session as the fMRI data for participants who did not already have a structural MRI scan available.

TMS

A 2x3x3 repeated measures design was used during the main TMS experiment. Experimental factors were AM rate (2 levels; 4Hz, 40Hz), TMS (3 levels; sham, left, right) and timepoint (pre-TMS, post-TMS-1 and post-TMS-2). The dependent variable was accuracy on the AM depth discrimination task. All subjects completed six TMS sessions in total; one for each permutation of AM rate and TMS condition. The order of sessions was approximately counterbalanced across subjects using a latin square design. At least 24 hours elapsed between each TMS session. A Brainsight frameless stereotaxic system (Rogue Research, Montreal, Canada) with chin rest and forehead support was used to monitor coil placement and ensure accurate TMS delivery. A Magstim Super Rapid^2 (Magstim, Whitland, UK) stimulator and a 80mm figure-of-eight TMS coil (Magstim, Whitland, UK; 80mm diameter of each coil wing) was used for pulse delivery. The TMS coil was always oriented with the handle pointing horizontally in an anterior direction, parallel with the midline. Stimulation consisted of a continuous theta-burst protocol; with bursts of 3 pulses separated by 5 ms delivered at a rate of 20 Hz. The pulse train was 40 seconds long giving a total of 600 pulses, in accordance with previous studies and the standard safety guidelines for cTBS (Huang et al. 2005; Rossi et al. 2009; Oberman et al. 2011). TMS pulse intensity was set to 40% of the maximum stimulator output (maximum = 2.6 T), which corresponds to $\sim$70 – 80 % of active motor threshold (Bestmann et al. 2003) and participants wore earplugs during cTBS. Before the first experimental TMS session, a short train of test pulses were delivered to ensure that subjects could tolerate this TMS protocol. The left and right TMS targets were obtained using the fMRI data as previously described. Sham TMS was performed by placing the coil at the vertex and rotating by 90°to remove any physiological effect of TMS but, crucially, retain peripheral factors such as the auditory artefact (Lisanby et al. 2001).

All auditory testing was completed in a double-walled sound attenuating booth. The design of the 2AFC task was identical to that used during baseline psychophysical testing, aside from the AM depth of the comparison interval was fixed (at 70% threshold) and participant feedback was removed. One block of 200 trials was completed immediately prior to the TMS train. Another block of 200 trials was then completed. The final block was completed 20 minutes after the previous block began.

Key Methodological Changes

This section will detail the key similarities and differences between the design previously described in Chapter 4 and the design used for this study. The main difference in the design of the behavioural task between studies described previously and the study described here is that an AM discrimination task was used in this study. In the AM discrimination task, 3 continuous 0.5 s noise periods are presented and period 1 and 3 are both amplitude modulated, however one AM period is modulated to a greater depth. This is different from the AM detection task used in previous chapters where only one noise period was amplitude modulated. An AM discrimination task was used here as AM discrimination more closely resembles the processes underlying AM processing during speech than AM detection.

The fMRI design used in this study was identical to the fMRI design reported in previous chapters.

The TMS protocol used in this study was continuous theta burst stimulation (cTBS). This is distinct from the TMS protocols used in previous chapters as cTBS is an offline, repetitive TMS protocol. Further, many TMS design decisions in the previous chapter were due to the need to recreate the conditions required to observe a state-dependent TMS effect. cTBS is a more established TMS protocol with a commonly used set of parameters, so this study followed these standards. The main advantage of cTBS compared to other repetitive TMS protocols is that efficiency is increased due to the large number of TMS pulses delivered in a short time. A further advantage is that the TMS artefact is not present during auditory presentation as cTBS is offline - stimulation is delivered then the behavioural task is completed during the aftereffect period. These reasons motivated the choice to use cTBS to investigate cortical AM processing.

Statistical Analysis

Generalised linear mixed models were fit to the data to examine statistical significance. This method is recommended over repeated measures ANOVAs in within-subjects designs as the hierarchical nature of the data generation is taken into account (i.e., explicitly accounting for multiple measures from the same participant leads to more accurate estimates) (Baayen et al. 2008; Dixon 2008; Jaeger 2008). Further, the binomial distribution of the raw data is problematic for standard linear modelling techniques, but the GLMM allows different link functions to account for this; a logit link function was used here. The dataset was subsetted based on AM rate and two separate models were fit due to a lack of convergence when AM rate and its interaction terms were included as fixed effects. Factors of TMS site and timepoint were modelled as fixed effects, and participant was modelled as a random effect. The random effects structure was the maximal justified by the design (Barr et al. 2013); the model with maximal random effect structure was fit then systematically reduced until convergence. Treatment coding was used to examine the difference from the baseline condition of pre-TMS timepoint (prior to left hemisphere TMS) and the post-TMS timepoints (separately). Finally, models were parametrically bootstrapped and 95% confidence intervals for all main-effect and interaction parameters were obtained. All modelling was conducted using the *lme4* package (Bates et al. 2015) in *R* software (R Core Team 2017). Parametric bootstrapping was conducted using the *pbkrtest* package (Halekoh and Hojsgaard 2014).

5.4 Results

5.4.1 Psychophysics

The modulation depth at which 70% accuracy was achieved on the AM depth discrimination task was calculated individually for each participant, for each AM test rate (4 and 40 Hz). The mean 70% threshold modulation depth across participants for the 4 Hz condition was -21.2 dB and for the 40 Hz condition was -24.4 dB. This is consistent with previous measures of the TMTF for AM depth discrimination of noise carriers (Wakefield and Viemeister 1990). A paired t-test showed a statistically significant difference between thresholds for 4 Hz and 40

Hz AM depth discrimination (t(9) = 2.58, p = 0.03).

Table 5.1 shows the 70% threshold values obtained from the behavioural runs for the AM depth discrimination task at 4 and 40Hz for each participant. These depth values were used for each trial in the TMS runs.

Participant	4 Hz threshold (dB)	40 Hz threshold (dB)
P1	-25.0	-28.3
P2	-25.3	-23.5
P3	-21.4	-31.1
P4	-17.4	-17.7
P5	-18.3	-18.8
P6	-23.3	-27.7
P7	-21.7	-21.6
P8	-17.5	-27.3
P9	-19.6	-21.6
P10	-22.0	-26.3

Table 5.1: AM depth discrimination 70% threshold values in dB for each participant. These values were used in the main task runs. The standard interval was always modulated at a depth of 0.36m. dB values are computed as $20\log(m)$.

5.4.2 fMRI derivation of TMS Targets

TMS targets in MNI-152 co-ordinate space for all participants can be found in Table 5.2.

Figure 5.2 shows targets in both hemispheres for an example participant.

TMS target distance measurements were recorded for TMS pulses. One participant lacked this data due to a technical issue. This is a measure of the distance from the centre of the TMS coil to the fMRI-defined target. The mean distance from the TMS coil to the left hemisphere target was 30.8 mm (SD: 6.889). The mean distance from the TMS coil to the right hemisphere target was 30.6 mm (SD: 8.449). There was a negligible mean difference of 0.212 mm between left and right hemisphere conditions.

TMS positioning error measurements were also recorded. One participant lacked this data due to a technical issue. This is a measure of the lateral distance

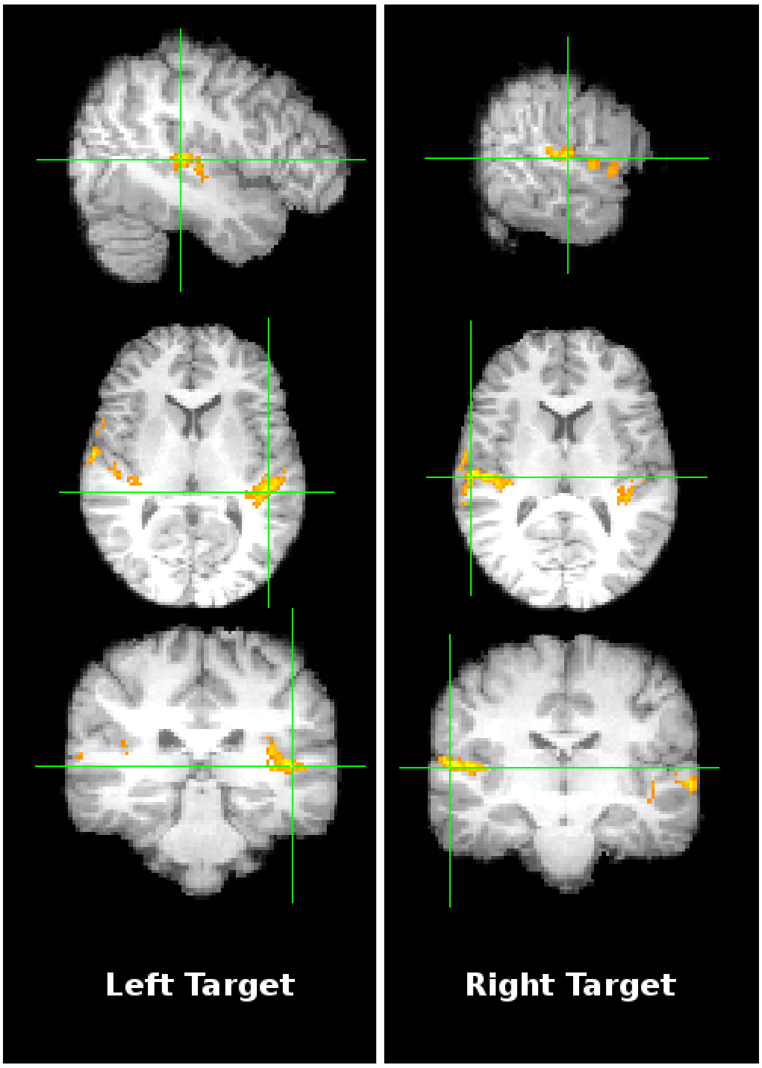

Figure 5.2: fMRI TMS target localisation data from an example participant. Overlay shows clusters of activation (Z >2.3; p <0.05) in bilateral pSTG/PT in response to a combination of 4 and 40 Hz sinusoidally amplitude modulated broadband noise conditions (combined) over unmodulated noise. TMS targets are indicated by the crosshairs. Targets were computed in individual coordinate space then transformed into MNI-152 co-ordinate space for the group analysis, and back transformed to participant co-ordinate space before targeting. Target locations for this participant in MNI-152 coordinates are: Left: $-46, -32, 6$ and Right: $62, -20, 10$

Participant	Left hemisphere target	Right hemisphere target
1	-64, -22, 8	60, -20, 8
2	-52, -26, 4	54, -28, 10
3	-50, -26, 0	54, -34, 10
4	-50, -36, 4	44, -26, 0
5	-60, -18, 0	62, -20, 10
6	-40, -32, 8	50, -24, 6
7	-58, -38, 14	42, -28, 8
8	-46, -32, 6	62, -20, 10
9	-44, -32, 10	50, -28, 0
10	-60, -34, 16	50, -28, 8
Mean	-52.4, -29.6, 7	52.8, -26.4, 6.6

Table 5.2: MNI-152 co-ordinates for TMS targets in left and right hemispheres for all participants. The average co-ordinate across all participants is also given.

from the fMRI-defined target to the estimated linear pulse trajectory. The overall mean error was 0.379 mm. The mean error in left hemisphere conditions was 0.381 mm (SD: 0.295) The mean error in right hemisphere conditions was 0.389 mm (SD: 0.341) There was a negligible mean difference of 0.008 mm between left hemisphere and right hemisphere conditions.

5.4.3 Behavioural Analysis

TMS Condition	AM Rate (Hz)	Timepoint		
		Pre	Post –1	Post –2
Sham	4	66.6%	64.1%	65.6%
Left	4	69.0%	69.2%	67.0%
Right	4	69.0%	67.8%	67.8%
Sham	40	70.4%	73.2%	70.3%
Left	40	72.6%	70.5%	72.0%
Right	40	70.0%	69.2%	72.3%

Table 5.3: Mean AM depth discrimination accuracy in % correct, averaged over all participants.

Mean accuracy was calculated for all conditions and can be found in table 5.3. A graphical representation of this data can be found in Figure 5.3.

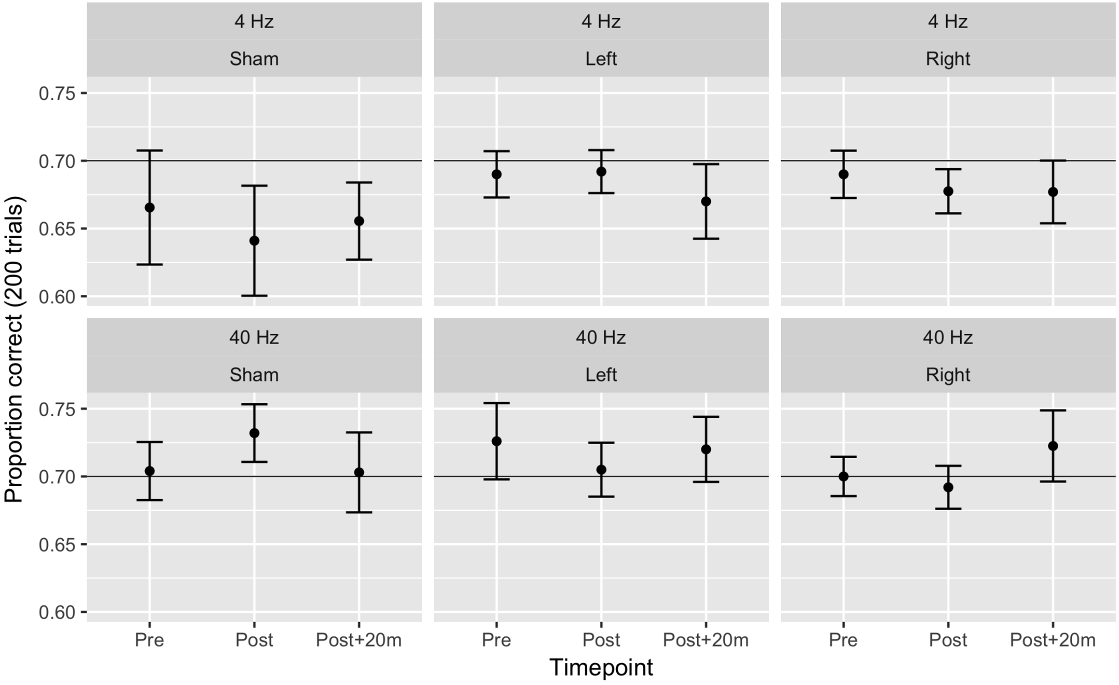

Figure 5.3: Mean accuracy across all participants for all experimental conditions. The horizontal line represents 70% accuracy. The upper row of subplots show data from the 4 Hz task while the lower row shows data from the 40 Hz task. The first column of subplots shows data from the sham TMS conditions, the centre column shows data from the left hemisphere TMS conditions and the final column shows data from the right hemisphere TMS conditions. Error bars show the 95% confidence interval.

5.4.4 Sham Summaries

For the sham 4 Hz AM condition, mean AM depth discrimination accuracy across all timepoints was 65.4%. For the sham 40 Hz AM condition, mean AM depth discrimination accuracy across all timepoints was 71.3%. For the 4 Hz AM condition, mean pre-TMS AM depth discrimination accuracy was 68.2%. For the 40 Hz AM condition, mean pre-TMS AM depth discrimination accuracy was 71.0% These summaries for the sham TMS and pre-TMS baseline conditions suggest that the 70% threshold estimation procedure worked well. It is also notable that the 4 Hz and 40 Hz scores are consistent across the separate control conditions, suggesting that sham TMS had little measurable effect on AM depth discrimination accuracy. While this increases confidence in using sham TMS as a control condition for the main test scores, the accuracy scores across timepoints within the sham TMS condition were highly variable; the largest variance of all conditions was found in the 4 Hz AM with sham TMS condition. This inconsistency means that using the sham TMS conditions as a baseline for statistical comparisons may be problematic. It is unclear why this pattern of high variance was found in the sham TMS conditions.

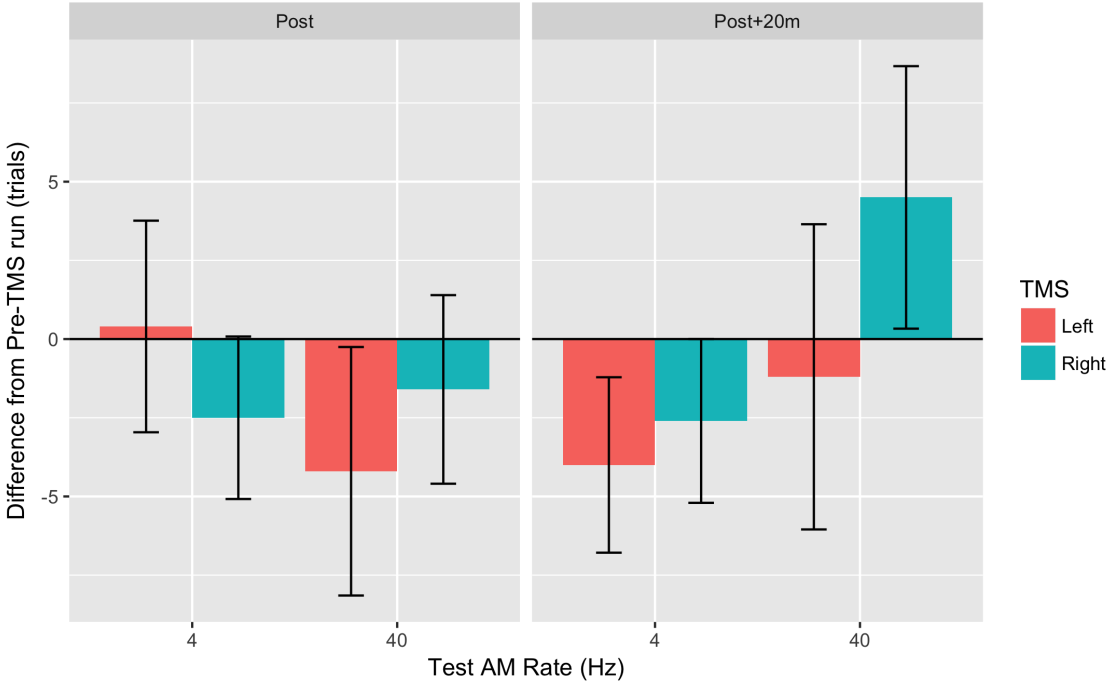

Figure 5.4: Changes in AM depth discrimination accuracy between the pre-TMS condition and left and right TMS conditions calculated as the mean of within-subject differences for each participant. The left hand graph shows results for the post-TMS timepoint 1 and the right hand side results for the post-TMS timepoint2 (20 minutes later). Both represent the difference from the pre-TMS baseline. Error bars show the 95% confidence interval.

5.4.5 Main Analysis

GLMM analysis of 4 Hz data

Results from this model are shown in Table 5.4. Factors included in this model are TMS condition (3 levels: sham, left hemisphere, right hemisphere) and Timepoint (3 levels: pre-TMS, post-TMS 1, post-TMS 2). Model fit values for this GLMM are as follows: AIC = 485.4, BIC = 541.9 . No statistically significant effects of cTBS on AM depth discrimination were found in any of the experimental conditions.

	Dependent variable:	
	Accuracy	95% CI
Baseline	0.839	(0.558 1.129)
Right Hemisphere TMS	−0.009	(−0.198 0.175)
Post-TMS Timepoint 1	0.007	(−0.157 0.173)
Post-TMS Timepoint 2	−0.047	(−0.247 0.159)
Right Hemisphere TMS x Post-TMS Timepoint 1	−0.064	(−0.274 0.146)
Right Hemisphere TMS x Post-TMS Timepoint 2	0.012	(−0.221 0.248)

Table 5.4: GLMM results from the 4 Hz AM conditions. Accuracy parameters and 95% CIs are in logit units. 'Baseline' represents the baseline condition of pre-TMS timepoint in the left hemisphere. All coefficients represent the difference from this baseline condition.

GLMM analysis of 40 Hz data

Results from this model are shown in Table 5.5. Factors included in this model are TMS condition (3 levels: sham, left hemisphere, right hemisphere) and Timepoint (3 levels: pre-TMS, post-TMS 1, post-TMS 2). Model fit values for this GLMM are as follows: AIC = 496.1, BIC = 552.7 . No statistically significant effects of cTBS on AM depth discrimination were found in any of the experimental conditions.

	Dependent variable:		
	Accuracy	95% CI	
Baseline	1.061	(0.672	1.455)
Right Hemisphere TMS	−0.179	(−0.415	0.056)
Post-TMS Timepoint 1	−0.124	(−0.335	0.081)
Post-TMS Timepoint 2	−0.021	(−0.297	0.253)
Right Hemisphere TMS x Post-TMS Timepoint 1	0.100	(−0.148	0.347)
Right Hemisphere TMS x Post-TMS Timepoint 2	0.185	(−0.144	0.507)

Table 5.5: GLMM results from the 40 Hz AM conditions. Accuracy parameters and 95% CIs are in logit units. 'Baseline' represents the baseline condition of pre-TMS timepoint in the left hemisphere. All coefficients represent the difference from this baseline condition.

5.5 Discussion and Conclusions

5.5.1 Overview

The current study aimed to investigate the neural mechanisms underlying AM depth discrimination using offline fMRI-guided continuous theta-burst stimulation. AM depth discrimination ability was measured at rates of 4 and 40 Hz before theta-burst stimulation and at two timepoints post-stimulation. TMS was delivered to 3 sites in separate sessions: vertex (sham), left auditory cortex, and right auditory cortex. The efficacy of sham cTBS for use in auditory experiments was also tested. No evidence supporting any of the three hypotheses was found as no statistically significant effects were found in any of the experimental conditions.

5.5.2 Possible Explanations for the Lack of cTBS Effect

Despite previous results showing the effectiveness of TMS for modulating AM detection sensitivity, the current results showed no effect of cTBS on AM depth discrimination. Two possible explanations for the lack of observed effects, namely the differences in stimulation method and AM task between this study and previous work are discussed here. It is well known that the intensity of the magnetic field generated by TMS decreases significantly as the distance from the coil increases. (Stokes et al. 2007). In contrast to the previous study where online dual pulse TMS was delivered at an intensity of 65%, the cTBS protocol used in this study was at an intensity of 40%. Further, the relatively deep TMS targets and low intensity protocol could have combined to result in a low intensity of stimulation induced at the target, possibly contributing to the lack of measured effect. Another important factor to consider is the different frequency of TMS used by each study. The previous dual pulse TMS study used a inter-pulse-interval of 50 ms, resulting in two pulses of 20 Hz stimulation. In contrast, cTBS is comprised of a train of 3-pulse bursts at 50 Hz, each separated by 200 ms. As AM rate is a key factor in this study and the close alignment between frequencies of oscillatory brain activity and auditory processing at these rates has been demonstrated, it is possible that this differential frequency impacted on the result. Although each stimulation protocol was chosen in a principled way based on previous research, interactions between stimulation rate and AM rate of the auditory stimuli must also be considered. The ability of single pulse TMS to entrain oscillations at particular frequencies has been demonstrated (Thut et al. 2011), however there is evidence for cTBS having no effect on oscillatory activity (Müller et al. 2013). This is one avenue of further research as the true importance of oscillatory activity to auditory and AM processing continues to be elucidated. Either of these methodological differences (or a combination) could have resulted in the differential TMS effects observed between the two studies.

Further to the methodological differences in the TMS protocols, the studies also differed in the behavioural task used. Previous work showed that TMS could affect AM sensitivity as measured by accuracy on a simple AM detection task. The current work used a higher order AM depth discrimination task where more subtle judgements were made. This task was chosen as it more closely represents the process of AM extraction during speech processing (Schlittenlacher and Moore 2016). There are two primary reasons why this may have affected the results. Firstly, although TMS targets were calculated from a contrast of AM noise over unmodulated noise, it is possible that the extraction of this higher order information (namely the difference in AM depth between two signals) is actually conducted in a different brain location, or even spread across multiple locations in stages, as AM filterbank models may suggest. If this is the case, it may be that TMS to a single location would not suffice to disrupt this form of AM processing. Secondly, a key distinction in this study compared to previous studies investigating motor activation over time is that rather than participants passively sitting while EMG measurements were taken at specific timepoints, participants in this study were actively completing the behavioural AM task during much of the 20 minute period post-cTBS. It is possible that this continual activation interacted with the cTBS effect and diminished it. Exploring this possibility further would be an interesting line of enquiry for future research.

Although no effects reached significance in the GLMMs, is it interesting to note that a trend observed in the data measured 20 minutes post-TBS does conform to predictions based on a combination of previous studies. Andoh and Zatorre (2013) measured activity increases in contralateral auditory cortical regions after cTBS and found that right hemisphere cTBS resulted in increased contralateral activity whereas left hemisphere cTBS did not. This was also reflected in behavioural measures as reaction times were graded such that a higher facilitatory effect was accompanied by a larger BOLD response in left anterolateral Heschl's gyrus. Following the predictions of the Asymmetric Sampling in Time hypothesis (Poeppel 2003), if the left hemisphere is specialised for processing signals with higher rates of change, increased activity in left auditory areas could lead to specific increases in sensitivity to these signals. A subtle increase in 40 Hz AM sensitivity is shown with right hemisphere stimulation, however a small decrease of 4 Hz AM sensitivity with left hemisphere cTBS and possibly also right hemisphere TMS was also shown. This dissociation could be explained by a combination of the effect previously observed by Andoh and Zatorre (2013) and predictions following the Asymmetric Sampling in Time hypothesis (Poeppel 2003). Further investigations are required before more conclusive statements could be made, but this initial study provides a specific timepoint and hemispheric balance of effects for future work to focus on. A further focus for future work could be on identifying and reducing sources of variance in order to improve the reliability of cTBS generally and specifically for auditory applications.

5.5.3 High Inter-individual Variance

Particularly high variances were found for AM depth discrimination accuracy in the sham conditions. This was unexpected and could be improved by increasing the number or modifying the design of behavioural task runs participants completed before the main task began. Variability has been shown in other temporal resolution tasks (Smith et al. 2008: e.g. gap detection:), although this has high test-retest reliability after 4 sessions. In fact, much of the psychophysical literature includes many hours of task practice before the study begins (Wakefield and Viemeister 1990: e.g.) and psychometric functions are assumed to be stationary, however for multimodal studies, this extensive pre-testing is often unfeasible and this assumption does not hold. A more feasible solution may be to improve the psychophysical methodology in order to account for non-stationarity (Doll et al. 2015). The unexpected variance observed in the sham conditions are likely due to comparatively naive participants and non-stationary observers. This is not fully addressed in the auditory psychophysics literature but more recent studies have implicitly accounted for this, for example, Schlittenlacher and Moore (2016) added an easily detectable reminder stimulus every five trials. Future work should account for this aspect more explicitly and recent studies have begun to address the need for efficient adaptive staircasing and function fitting methods that are suitable for use with naive or non-stationary participants (Shen 2013; Shen and Richards 2013; Shen et al. 2015; Rinderknecht et al. 2018).

High inter-individual variability is often observed with non-invasive brain stimulation methods and this also holds true for cTBS (López-Alonso et al. 2014). Previous studies investigating TBS effects on auditory function have also found this (Lorenz et al. 2010; Müller et al. 2013; Andoh et al. 2015). This high variability is thought to be caused by many variables including genetic factors, age, variations in intracortical network activity and functional connectivity (Suppa et al. 2016). This variability, as with many other methodological factors relating to cTBS has been more thoroughly measured using motor areas (Vernet et al. 2014), but areas with less well characterised behavioural effects are more difficult to assess. Perfusion MRI is one possible technique that could be used to measure this in the future as it has been found that cerebral blood flow is a marker of inter-individual variability in responses to TBS (Gratton et al. 2014). Specifically, this study found that changes in functional connectivity were directly related to changes in cerebral blood flow due to TBS for specifically targeted brain networks (and not for a control site). Functional connectivity changes were also measured when TBS was delivered to auditory areas (Andoh et al. 2015), suggesting a possible measure for indexing this variability in the auditory domain. In this case, greater anatomical connectivity measured in the transcallosal auditory pathway was directly linked to greater TMS induced changes in inter-hemispheric functional connectivity. Reducing this inter-individual variability is important for group averaging techniques and high levels of variance make the generalisability of results beyond the observed sampled increasingly difficult to support. Future work could ameliorate this issue by more extensively measuring individual difference profiles in small samples, behavioural and connectivity correlates.

5.5.4 Future Considerations

Key variables to control in order to minimise this variability can be identified from TBS and behavioural auditory experiments. Accounting for some behavioural factors may help reduce variability. The shape and cutoff point of the temporal modulation transfer function (TMTF) is known to vary inter-individually (Viemeister 1979). If 4 or 40 Hz AM thresholds differ between participants, it suggests that the shape of their TMTFs differed. To further investigate the possible influence of this factor, full TMTFs could be measured and the effect of

TMS on AM detection at different rates compared across TMTF shapes as a whole, rather than focussing on specific rates. Efficient methods of estimating the two TMTF parameters have been recently developed, making this approach more feasible (Shen et al. 2015). Recreational noise exposure is also known to affect AM sensitivity even when hearing thresholds are within normal limits (Stone et al. 2008; Kumar et al. 2012). Deficits that do not manifest in the audiogram results are known as hidden hearing loss and are thought to result from permanent damage to synaptic connections between inner hair cells and auditory nerve fibres. However, this is not fully understood and paradoxically, damage due to recreational noise exposure can lead to an increase in sensitivity to AM (Stone and Moore 2014). Understanding hidden hearing loss and improving current methods of measuring it is an active research topic (Dewey et al. 2018; Guest et al. 2018: e.g.). Using newly developed measures and investigating how they relate to AM sensitivity would be greatly beneficial to our understanding of AM processing in normal hearing and hearing impaired individuals. Until such techniques are developed, taking measures of TMTF shape and recreational noise exposure in order to control for these factors beyond the audiogram would be a feasible way of examining and explaining behavioural variance.

In addition to modifying the behavioural task, the aforementioned measures of structural and functional connectivity could be taken and correlated with TBS response. The degree of processing laterality could also be an important factor, especially in studies where auditory hemispheric asymmetry is a key variable. The laterality of auditory processing is known to vary between individuals and multiple electrophysiological and functional neuroimaging measures have been developed to measure it. Measures such as the acoustic change complex (Han and Dimitrijevic 2015), the right ear advantage (Hugdahl 2011) or simply the average asymmetry of activation in auditory areas from an fMRI localiser could be correlated with response to TBS. Further, current modelling tools that integrate structural connectivity information (Geeter et al. 2016) may be able to model these asymmetrical effects in future, when the connectivity profile of auditory areas is characterised to a greater degree. Fully exploring the influence of these extra sources of variance would advance our understanding of cTBS effects in auditory experiments and inform future studies of the critical factors to control in order to reduce variability.

5.5.5 Conclusions

This study was designed to further investigate whether cTBS to auditory cortical regions can be used to affect AM processing and if the effects align with asymmetric accounts of auditory processing. From these data, cTBS was not found to be an effective method for modulating AM depth discrimination ability. Possible explanations for this results were discussed, primarily that the intensity of the standard cTBS protocol may not be high enough to induce significant electrical activity in the relatively deep targets localised from the fMRI contrast. Further aspects to be considered include the frequency of magnetic and auditory stimulation, possible asymmetrical network effects and controlling for the many sources of between-subject variance in auditory behavioural tasks and cTBS. While no cTBS effects were measured on AM depth discrimination, this is not evidence for the absence of a true underlying effect or that all TMS protocols are equally ineffective. Further research is required to determine the effectiveness of cTBS for modulating AM sensitivity, although methodological changes are recommended.

Chapter 6

General Discussion

6.1 Overview

This investigation was designed to address the following three overarching research questions:

- Firstly, is fMRI-guided TMS an effective method for modulating AM processing?
- Secondly, are different TMS protocols more or less effective at modulating AM processing?
- Finally, is fMRI-guided TMS an effective method for further understanding the functional asymmetry of speech processing?

Three extensive multimodal experiments were conducted and the results demonstrated the following. Firstly, it was shown that online double-pulse transcranial magnetic stimulation (TMS) is an effective tool for modulating AM sensitivity. Before this work, it was not clear whether TMS was able to perturb processing of this low-level aspect of speech at all. Utilising this fairly novel tool (Barker et al. 1985), new causal evidence was found for the contribution of posterior superior temporal gyrus (pSTG) / planum temporale (PT) to AM processing. Specifically, it was shown that fMRI-guided online TMS perturbed detection of 4 Hz AM when delivered to the right hemisphere only. This finding was interpreted using contemporary models of asymmetric auditory processing, however, as only one hemisphere showed an effect, conclusions about asymmetry were difficult to support or refute.

Secondly, it was shown that double-pulse TMS can also be used to modulate AM sensitivity in a state-dependent way. Auditory behavioural adaptation was used to manipulate the underlying state of the cortex when TMS was applied. TMS to the left hemisphere affected 40 Hz AM sensitivity, however a dissociation was observed where adaptation to 40 Hz AM led to an disruption of AM sensitivity, whereas adaptation to 4 Hz AM led to a facilitation.

Finally, it was shown that continuous theta burst stimulation (cTBS), an offline, repetitive TMS protocol, was ineffective at modulating AM sensitivity. Accuracy was measured on a higher order AM task that more closely represents the processes involved in parsing speech. The inability of cTBS to affect this task may be due to the low stimulation intensity and relatively deep cortical targets or it might be a result of high inter-individual variability of the chosen methods. These considerations were discussed in depth and suggestions were made for how future studies could reduce sources of variance, relating to both auditory behavioural tasks and cTBS.

A summary of the primary findings from this investigation, directly related to the initial research questions, are as follows:

- fMRI-guided TMS can be an effective tool for investigating neural processing of AM. Building on this initial finding, it is now possible to design future studies incorporating TMS to further investigate many aspects of low-level auditory processing.

- Multiple fMRI-guided TMS protocols can be used to modulate AM processing, and their effectiveness does vary. Online double-pulse TMS and a TMS-adaptation paradigm were shown to be effective though an offline continuous theta-burst paradigm was less effective.

- fMRI-guided TMS can be an effective method for further understanding the functional asymmetry of speech processing, however, some caveats must be accounted for. These studies were developed in order to test asymmetric models of auditory processing, primarily the Asymmetric Sampling in Time (AST) hypothesis (Poeppel 2003) and where an effect was found, its relevance to asymmetrical auditory processing frameworks was discussed. However, as absence of evidence is not equal to evidence of absence, where an effect was found for only one hemisphere or AM rate, interpreting this outcome is difficult. This is a key consideration for future studies that aim to investigate TMS effects on functional asymmetries and is discussed further.

Recent methodological developments in psychophysics, fMRI and TMS could be incorporated to substantially improve future research.

6.2 Psychophysics

Designing auditory psychophysical tasks, fitting psychometric functions and extracting thresholds constituted a significant part of the studies described here. Estimating a single threshold is a complex procedure that consists of multiple experimental stages containing many variables that can each affect the final threshold value. Many of these variables would benefit from further study in order to fully optimise the procedure. This is true of psychophysical methods generally but particularly for measures of temporal resolution and sensitivity to AM. It is known that temporal processing tasks are affected by both sensory and non-sensory factors, such as memory (Smith et al. 2008), but there is a paucity of further research into this area.

Task design details such as the number of intervals presented and the temporal order of intervals can have a large impact on resulting parameter estimates. The number of alternative intervals presented in each trial primarily affects the baseline chance level but is also of particular relevance for temporal designs where intervals are not presented concurrently. It is not known what additional demands these designs have and how this affects interpretation of the results. The AM detection tasks used in these studies were primarily based on those used by Viemeister (1979) and the AM discrimination task was based on Wakefield and Viemeister (1990), however, piloting showed that the reliability of thresholds could be increased by adding a 500 ms noise period to the start and end of the trial. The choice of task design can also give rise to biases and dissociating the perceptual effects from response or decision biases is crucial to determining the true effect size and psychometric function parameters (Morgan et al. 2013; García-Pérez and Alcalá-Quintana 2013). An example of a response bias is the Type B effect, where participants are more likely to identify the interval presented second as the target in a temporal 2AFC task (Ulrich and Vorberg 2009; Dyjas and Ulrich 2014; Bausenhart et al. 2015). The extent to which AM detection and depth discrimination tasks are affected by these biases is not fully explored, although similar tasks have been examined (auditory duration discrimination) (Lapid et al. 2008; Rammsayer and Ulrich 2012). Fully characterising these effects in auditory AM detection and discrimination studies would minimise response biases and allow the design of more specific tasks.

Further variability arises due to differences in staircase design and the choice of optimisation algorithm used. The simplest adaptive method uses up/down rules and systematically decreasing step sizes to estimate the psychometric function by honing in on a specific threshold value (Levitt 1971). The method used here involved fitting the full psychometric function and then extracting the predicted threshold value. An alternative approach involves using a set of fixed step sizes and then taking an average of the last reversal values to directly obtain the threshold value (but no visual representation of the psychometric function). However, there are more advanced adaptive staircasing methods (Watson and Pelli 1983; Treutwein 1995) and new approaches are still being developed (Schütt et al. 2016; Bak and Pillow 2018). These methods improve efficiency by using optimisation techniques to adaptively modify the step size in order to maximise the predictive value of each trial. Of particular note is a recently developed method that was specifically designed for efficient psychometric function estimation and validated with an auditory temporal resolution task (Shen and Richards 2012; Shen et al. 2015). Implementing this updated maximum likelihood procedure in future experiments would improve efficiency and enable more accurate threshold estimates.

As part of the threshold estimation for these studies, multiple task runs were conducted and separate functions were fit for each run (though multiple interleaved staircases were used in each). The function with the highest threshold value was discarded in order to remove the influence of poorly fit functions, typically resulting from early errors which can be difficult for the staircase to recover from as step sizes are large. An average of the remaining threshold values was then calculated to provide the final threshold value. Whilst this method of choosing which functions to include is principled and consistent, it is possibly flawed. The choice could be improved by using goodness of fit metrics, rather than implicitly assuming the highest threshold is the poorest fit. Also, combining data from the remaining runs and fitting a single function may increase accuracy, however this may underestimate the slope value if thresholds are non-stationary (Wallis et al. 2013).

The choice of threshold value to target during the threshold estimation stage is important. Different variances have been measured in the same listeners when the target threshold is varied and the most accurate psychometric functions are measured when a ‘sweet point’ on the threshold is targeted. In order to measure the threshold with minimum expected variance, targeting a higher threshold than the standard 50% or 70.7% has been recommended (Green 1990). The design of the adaptive staircase, including the choice of specific step sizes and reversal rule determines which threshold percentage is targetted. This is further complicated when lapses of attention are considered. Lapse rate represents the probability of a correct response when the stimulus was not seen, e.g. 50% for a 2AFC task. Including a parameter for lapse rate in the psychometric function is important as it has been shown that other parameter estimates may be biased if lapse rate is constrained to 0 (Wichmann and Hill 2001; Prins 2012). The aforementioned updated maximum likelihood procedure targets the sweet points of multiple parameters, and has been shown to be effective in the case of a high lapse rate (Shen and Richards 2012). Ideally, a widely used, well-characterised 'standard' psychophysical task would be used, but it is important to remember that many aspects of study design are constrained by the inclusion of TMS. The timing of TMS delivery in particular is a major consideration when looking to combine these two methods. For a more nuanced example, the choice of (fixed) threshold value used for the main tasks was guided by prior knowledge of TMS effect variability. Taking this into account, thresholds were chosen in these studies that were reasonably resistant to ceiling and floor effects in either direction. Now initial evidence for an online TMS disruption to AM detection has been found, it would be interesting to systematically exploring this factor. Future research could measure the full psychometric function post-TMS and quantify the effect in more detail by investigating which parameters are most affected and how the deficit manifests.

When these factors are accounted for, a major issue still remains if thresholds are non-stationary (Schütt et al. 2016). In psychophysical research it is very common to use practised participants and assume a stationary psychometric function. However, if the threshold is not stable across or within testing sessions, estimation becomes more complex. In many previous studies, participants have had multiple hours of practice before data collection begins (Viemeister 1979: e.g.). As multimodal brain imaging and stimulation studies commonly consisting of multiple extensive testing sessions, this approach may not be feasible. This issue has been addressed in the case of young participants (Witton et al. 2017) and other psychoacoustic tasks (Amitay et al. 2006), but not for AM detection and depth discrimination thresholds with adult participants. Further understanding how non-stationarity affect AM detection and depth discrimination thresholds is important for three reasons. Firstly, fewer behavioural training sessions would be required. Secondly, better estimates of the underlying variance would lead to more accurate determinations of the sample size required to reliably estimate effects Finally, a broader question exists regarding how generalisable the results from experienced participants are to the general population. If participants undergo hours of training and their threshold stabilises, it is possible that they are using a psychological strategy or their auditory system has adjusted in some way to improve performance. In this case, it is difficult to support that the processes involved in the task are identical to those used under normal listening conditions. One possibility is that aspects other than AM are utilised for speech processing so that AM extraction does not have to be developed to the particularly fine level reached through hours of behavioural task practice. However, this broad criticism is not specific to measures of AM sensitivity and could be equally applied to at-threshold procedures generally. Further understanding the processes by which participants learn to improve at the task and how their thresholds stabilise would lead to the development more efficient psychophysical methods and ore generalisable results.

To summarise, many factors are involved in estimating the threshold value and the effect of many of them are yet to be fully explored for AM tasks. It is not fully known how the temporal design of the AM tasks affects threshold estimation or if this introduces any biases. Particular psychophysical aspects that would benefit from careful consideration include optimal trial design, optimal staircase design and psychometric function form. Some design constraints are applied when brain imaging or stimulation is combined with psychophysical behavioural tasks. For example, task design will depend on stimulation timing and threshold non-stationarity is more likely to be a concern as intensive task practice may not be feasible when the behavioural session is just one part of a substantial multimodal experimental procedure.

6.3 TMS methods

The application of non-invasive brain stimulation techniques to investigate human cognitive function was developed relatively recently and methodological advances are frequent. Since this work began, significant developments have been made and similar investigations could be improved by incorporating some of them. Recent progress in the area of non-invasive brain stimulation current flow modelling have led to improvements of model accuracy and software implementations that allow the technique to be more widely adopted (Geeter et al. 2016; Thielscher et al. 2015). When conducting current flow modelling, researchers create anatomically accurate head models from structural MRI scans that include segmented layers with different conductive properties (Thielscher 2013; Goodwin and Butson 2015). It is then possible to model the expected current flow within a particular brain area using conductivity profiles calculated from physiological studies and individual brain topography. These techniques provide an additional source of evidence for the putative effect of brain stimulation and have been shown to improve accuracy of stimulation compared to simpler models. Future studies could be improved by the inclusion of current flow modelling to estimate the effective current density induced at the target location in auditory cortex. This information could be used to optimise TMS coil placement or in this case, could have been used to further investigate the lack of TBS effects. It would be particularly beneficial to examine the effect of gyrus orientation on induced current at an individual participant level due to the high inter-individual variability of auditory areas. Further, this would help inform the feasibility of new TMS investigations of auditory processing as gyrus orientation on the temporal plane are unlike that of M1, where most TMS methods are initially developed. Improving the accuracy of individual models of current flow will increase the validity of the non-invasive brain stimulation methods and inform future studies.

Whilst incorporating these modelling techniques would provide further insight into the immediate spread of locally induced current, accounting for more widespread network effects is more challenging. The assumption that TMS to a cortical area has an isolated local effect is naïve and the conceptualisation of TMS as solely creating virtual lesions has been superseded by more nuanced views (Ruff et al. 2009; Silvanto and Cattaneo 2017). Evidence has been found for widespread network interactions (Bestmann et al. 2004; Fox et al. 2014; Hallam et al. 2016), but the mechanisms underlying these effects are not fully understood. Network effects have also been found with TMS to auditory areas and these effects were shown to be asymmetrical (Andoh and Zatorre 2013; Andoh et al. 2015). Andoh et al. (2015) showed that cTBS to the right, but not left, hemisphere led to widespread decreased functional connectivity in the consequent resting state period. Whether the decrease in connectivity following cTBS in auditory regions remains during a behavioural task is an interesting question and the effect of different TMS protocols is also not known. It has also been shown that network changes due to TMS do not necessarily manifest as behavioural effects (Hallam et al. 2016). It is likely that network effects are present, despite no modulation of AM sensitivity being measured behaviourally, as even subthreshold TMS can lead to network interactions (Bestmann et al. 2004). Even with 'simple' online TMS effects, such as those shown in chapter 3, the immediate effects of TMS at the target and wider network interactions are rarely explicitly dissociated. Accounting for these effects is increasingly important as our understanding of the connectivity profile of the brain expands. For example, the simplistic assumption of a direct mapping between a targeted anatomical region and a specific function does not always hold (Margulies and Petrides 2013; Opitz et al. 2016). This is further complicated by the finding that the frequency of stimulation can lead to qualitatively different effects on functional connectivity (Eldaief et al. 2011), even within well-characterised motor networks where a direct mapping is more often assumed (Rounis et al. 2005). Broad network effects represent a potential source of variance that could lead to non-linear and unpredictable effects, especially with cTBS protocols. The existence of complex widespread interactions is well established, but current explanations of TMS effects are not sufficient to explain them (Silvanto and Cattaneo 2017: e.g.). In addition to the asymmetries in functional connectivity measured by Andoh et al. (2015), structural connectivity networks have also been found to be asymmetrical (Mišić et al. 2018). Integrating structural connectivity information into current flow models has been shown to improve modelling accuracy (Geeter et al. 2016). Further, inter-individual variation in structural connectivity is likely to account for some of the inter-individual variance in TMS effects so may utilising this information may lead to more accurate predictions of TMS effects or go some way to explaining differences between responders and non-responders. Targeting specific functional networks is a promising future option As the mechanisms underlying TMS effects remain largely unknown and further complexity is revealed, it is important that emphasis is placed on theory-grounded hypotheses and converging sources of evidence for the putative effects of brain stimulation. These studies show that structural and functional connectivity are likely to play an important role in mediating the effects of TMS, particularly in relation to network effects. Further multimodal studies are required to fully explain these network effects and new TMS studies should aim to integrate structural and functional connectivity information where available.

Determining the specificity of TMS effects is crucial and controlling for the extraneous effects of TMS is a challenging methodological problem (Adank et al. 2017). TMS pulses cause a loud auditory click and stimulate peripheral nerves incidentally, causing muscle twitching that varies significantly by stimulation location. Sham stimulation protocols have been developed to control for these factors, but they can be highly variable between studies. For example, the one-wing 90°vertex sham from Lisanby et al. (2001) was used here whereas other studies have used TMS to the shoulder (Herring et al. 2015). An ideal sham condition would recreate the exact sensory experience of active stimulation but this is only possible for low TMS output intensities and specific locations where no cutaneous stimulation effects are detectable. A common compromise is to deliver active stimulation to a location that is ostensibly not involved in the task of interest, however this is not ideal due to possible remote effects that are difficult to measure. A new tool allows twitching and participant annoyance to be matched between sham and active sites, which could improve sham stimulation (Meteyard and Holmes 2018). However, if multiple sites are of interest, including matched sham conditions can rapidly become unfeasible in many cases and a control task may be preferable (Adank et al. 2017).

The studies in this thesis were designed to investigate the existence of double dissociative effects between hemisphere of stimulation and AM rate. With this design, if no effect is found for one hemisphere or AM rate condition it can be difficult to dissociate a failure of TMS to cause an effect from a true inexistence of an underlying effect. In these initial investigations, where no previous examples of these effects exist, determining this difference is further complicated as conclusions are limited to the specific TMS and behavioural protocols used, yet researchers might be dissuaded from pursuing this line of investigation when initial results are inconclusive. However, now this initial evidence has been found, work can be conducted to determine the key parameters required to elicit an effect. In future investigations, a positive control task could be included to support the conclusion that TMS is affecting the target location in the absence of behavioural effects. Positive controls are dependent measures that are known to be affected by TMS in a predictable way and serve to demonstrate that TMS has affected the target region. Including a positive control task would help to determine whether, for example, deeper cortical regions were affected by TMS. For this to be feasible, a task that is reliably affected by TMS to the location of interest would need to be identified. Alternatively, future work could include AM detection as a positive control task when these auditory areas are being targeted by TMS. For example, it is known that frequency modulation is processed in similar auditory regions and may be represented orthogonally to periodicity in auditory cortex (Barton et al. 2012). Whether TMS can affect frequency modulation detection is an open question and one that can be addressed now this foundation work has been completed.

To summarise, the field of non-invasive brain stimulation is in its infancy and substantial developments in TMS methodology continue to be made. Future investigations could be improved by incorporating some of these advances. Current modelling techniques are particularly promising for targeting TMS and providing support for putative neural effects. Wider network effects of TMS are undeniably complex but need to be considered and incorporating structural or functional connectivity measures may help to explain variance of TMS effects. Appropriate and feasible sham TMS conditions are required to determine the specificity of any TMS effects and the inclusion of a positive control task would aid interpretation of future investigations where a full double dissociation is not found.

6.4 Alternative Approaches

It is important to consider findings from complementary approaches in order to retain a holistic view of speech processing and to understand how this work is positioned in the literature. The studies that comprised this project were primarily focussed on the perceptual consequences of brain stimulation, as measured by AM detection and discrimination tasks. Whilst the AM discrimination task approximates speech processing more closely than AM detection or gap detection tasks do, many other relevant cues are necessarily lacking, for example, speech-like spectral structure and higher level linguistic aspects. Whilst this bottom-up behavioural approach is valid and informative, a common alternative approach investigates the effect of magnetic, electrical or auditory stimulation by directly measuring neuronal activity. Complementary approaches that present the full speech signal, or a signal reduced in different ways are crucial for understanding the relative importance of each speech component (Hamilton and Huth 2018). Fully understanding all aspects of speech processing is beyond the scope of what any one approach can contribute and by examining the findings from other approaches, general principles underlying speech processing and knowledge gaps can be identified.

6.4.1 Neural Oscillations and Speech Processing

It is well established that speech is processed concurrently on multiple timescales (Peelle and Davis 2012; Giraud and Poeppel 2012). Evidence for this account has come from investigations using psychophysical (Chait et al. 2015), invasive electrophysiological (Lakatos et al. 2005), non-invasive electro- and neurophysiological (Gross et al. 2013; Ding et al. 2016; Teng et al. 2017), haemodynamic (Davis and Johnsrude 2003; Obleser et al. 2007), and computational modelling methodologies (Santoro et al. 2014). Initial accounts focussed on the high correspondence between repetition rates of key linguistic features and prominent neural oscillatory frequency bands (Giraud and Poeppel 2012). Of particular note are the theta and low gamma frequency bands as they closely correspond to the syllabic (∼4 Hz) and phonemic (∼40 Hz) rate of speech. It was proposed that this correspondence was a reflection of processing using different length temporal sampling windows (Poeppel 2003). Further, this processing was proposed to be asymmetric with each cerebral hemisphere exhibiting a preference for sampling windows of different lengths.

Since these early accounts, complementary explanations have emerged for the relevance of neural oscillations to processing of auditory stimuli. It is thought that these oscillations reflect cyclical phases of low and high excitability within neural populations and their oscillatory phase aligns with rhythmic input to facilitate processing (Bishop 1932; Schroeder and Lakatos 2009; Luo and Poeppel 2007). For example, temporal alignment of oscillatory activity in the theta band to the syllabic patterns of speech is thought to play a key role in synchronised processing (Hyafil et al. 2015). At this rate, high complexity syllables are followed by low complexity periods with an average cycle length of ∼250 ms. This phase entrainment to the speech envelope enables the complex components to be processed when neural excitability is highest. Oscillatory activity within other frequency bands have been found to entrain to other speech aspects and nested neural oscillations measured during presentation of speech stimuli are a possible substrate for an hierarchical framework that subserves synchronised linguistic processing

(Zoefel and VanRullen 2016; Gross et al. 2013; Ghitza 2017; Teng et al. 2017). Ding et al. (2016) measured the electrocorticographic neural response to quantised Chinese sentences and showed that 4 Hz activity aligned with word-level information (monosyllabic words in this case), 2 Hz activity aligned with phrase-level information and 1 Hz activity aligned with sentence-level activity. Crucially, this was shown to be dissociable from the acoustic content as the peaks at 1 Hz and 2 Hz were not present when the same stimuli were presented to non-Chinese speakers. This study reported these findings as evidence for a nested hierarchy of oscillatory activity, driven by speech linguistic content, however this is controversial as it is possible for a model with access to only lexical-level information to account for this neural data (Frank and Yang 2018). Further, the artificial nature of the stimuli confounds interpretation as natural speech is aperiodic and not perfectly quantised, though these oscillatory bands are still reflected in natural speech processing. Despite these limitations, this is just one example of emergent evidence for neuronal oscillations in auditory areas as a functional substrate for the discretisation and multiplexed processing of speech (Meyer 2017; Zoefel et al. 2018) Amplitude modulations of the speech envelope may have a crucial role to play in this system. Robust tracking of the speech envelope by human auditory areas has been found and temporal modulations in critical frequency bands are suspected to play a key role in speech processing (Kubanek et al. 2013; Ghitza 2011). This envelope tracking may facilitate processing by enabling phase resets of delta band activity, ensuring entrainment and efficiency (Doelling et al. 2014).However, slow amplitude modulations are just one piece of the system, and their well-characterised nature may have led to an overemphasis of their importance (Obleser et al. 2012). For just one example, evidence for entrainment to higher level acoustic features of speech suggests oscillations are not purely driven by low level acoustic features (Zoefel and VanRullen 2016).

It remains controversial whether synchronous entrainment actually plays a crucial role in speech processing or is purely epiphenomenal (Zoefel et al. 2018). A train of phase-locked responses evoked to repetitive rhythmic stimuli would be largely indistinguishable from intrinsic neural oscillations at the stimulus presentation frequency but it has been suggested that endogenous and exogenous oscillations may be functionally distinct (Meyer et al. 2018). Latest theories posit that this oscillatory activity measured by local field potentials or M/EEG may actually be better characterised as synchronised burst events (van Ede et al. 2018). Neural oscillations constitute a candidate mechanism for synchronised segmentation and processing of sensory inputs in multiple domains and are not specific to speech (Murphy 2015; Haegens and Zion Golumbic 2018; Ronconi et al. 2017). The notion that temporal sampling windows discretise continuous input is also not speech-specific and evidence for multiple key rates has also emerged in vision research(Ronconi et al. 2017; Holcombe 2009). Specifically, higher frequency gamma oscillations have been linked with finer sensitivity in both vision and audition and this is suggested to be a perceptual consequence of a higher sampling rate (Baltus and Herrmann 2015). These similarities across multiple modalities suggest that multiplexed neural coding may be a general underlying principle of continuous segmentation and integration of sensory information.

To summarise, there is a wealth of evidence linking neural oscillatory activity and speech processing. There is mounting evidence for a hierarchical processing architecture that with nested oscillations at key frequency bands. Theta and low gamma have been closely linked with speech processing, initially due to their close correspondence to rates of linguistic speech aspects. Auditory cortex is able to continuously track the amplitude envelope of stimuli and this has been exploited in studies of neural entrainment. Current theories posit that exogenous and endogenous oscillations may be functionally distinct and that some neural oscillations may be better characterised as synchronised bursting activity. Evidence for a hierarchical system of cascaded oscillations has been found for multiple sensory modalities, suggesting a general organisational system.

6.4.2 Non-Invasive Brain Stimulation and Neural Oscillations

Non-invasive brain stimulation research has been fundamental in testing these oscillatory accounts. These techniques have been applied to both disrupt ongoing oscillatory activity and ostensibly entrain oscillations at specific frequencies. Transcranial current stimulation (tCS) techniques in particular, have desirable properties for entraining oscillations and have been successfully applied to this end. This family of techniques includes transcranial direct current stimulation (Nitsche and Paulus 2000: tDCS:), transcranial alternating current stimulation (Antal and Paulus 2013: tACS:) and transcranial random noise stimulation (Terney et al. 2008: tRNS:) and involve passing a small current through the brain via electrodes placed directly on the scalp. The low intensity current (typically 1 or 2 mA) modulates the resting membrane potential of underlying neural tissue and has been shown to affect firing rate and timing. Despite their widespread use, mechanistic understanding of tCS effects has failed to keep pace with their myriad applications in both cognitive and clinical contexts (Bestmann et al. 2015). Another factor that has contributed to the abundance of tCS investigations into the role of neuronal oscillatory processing during auditory stimulation is their silent operation (Heimrath et al. 2016; Zoefel and Davis 2017). tCS techniques have been successfully applied to investigate and improve auditory temporal processing ability. Anodal tDCS to the left (but not right) auditory cortex was found to affect rapid temporal processing as measured by a gap detection task (Heimrath et al. 2014). 40 Hz (but not 6 Hz) tACS has been shown to disrupt perceptual learning in a phoneme categorisation task (Rufener et al. 2016). tRNS has been used to improve near-threshold gap detection performance (Rufener et al. 2017). The phase of 4 Hz transcranial alternating current stimulation has been shown to modulate sensitivity to click train stimuli (Riecke et al. 2015). Most promisingly for investigations of AM processing, temporal processing ability (measured by gap detection) has been enhanced using tACS with an AC rate above individual gamma frequency (Baltus et al. 2018). Multiple groups have also shown that incorporating the speech envelope into the electrical stimulation envelope can lead to increases in speech intelligibility (Riecke et al. 2017; Wilsch et al. 2018; Zoefel et al. 2018).

The putative explanation for these behavioural improvements with AC stimulation is active entrainment of neural oscillations at the stimulation frequency. It is well-supported that these oscillations are present during sensory processing and pre-stimulus sensory entrainment has been shown to lead to behavioural improvements (Ronconi and Melcher 2017; Haegens and Zion Golumbic 2018; Bosker and Ghitza 2018), but it is currently difficult to prove the existence of tACS entrainment. The most definitive evidence for active neural entrainment by tACS comes from direct measurement using concurrent M/EEG. However, this is controversial as the nonlinear interference created by the electrical stimulation is so destructive to the M/EEG signal (Noury et al. 2016; Neuling et al. 2017; Noury and Siegel 2018).

Multiple TMS approaches have been used to investigate the role of oscillatory activity on auditory processing (Pellicciari et al. 2017). TMS pulses have been shown to reset the phase of ongoing oscillations, with behavioural consequences that can be manipulated by adjusting the pulse timing (Herring et al. 2015; Thut et al. 2011). Oscillations can also be entrained by TMS, however, the method of effect is likely different from tCS (Thut et al. 2011; Albouy et al. 2017). TMS has some properties that enable investigations into the role of neuronal oscillatory activity in speech processing. In particular, exploiting the ability of TMS to deliver single pulses at specific timepoints is crucial (Thut et al. 2017). Utilising this in combination with EEG has allowed specific phases of oscillatory activity to be targeted (Pellicciari et al. 2017; Farzan et al. 2016). Combining this precise timing with direct readings of neural activity in a closed-loop, adaptive way is also a promising future direction (Price et al. 2015; Bergmann et al. 2016). Recent developments in understanding of the neural systems underlying neuroimaging and stimulation has allowed TMS to be applied in a more informed way with specific targeting of neural oscillatory phases or targeting of specific functional networks (Romei et al. 2016). Similarly to tCS, a more detailed understanding of the neurophysiological mechanisms underlying TMS effects is still inhibiting progress. These explanations would improve its feasibility as more specific theory-grounded hypotheses regarding the neurobiology of speech and language processing could be developed and tested. In this thesis, TMS was shown to be an effective modulator of AM sensitivity, primarily conceptualised as a disruptive interruption to ongoing processing in auditory cortex. Oscillatory phase resetting could be the mechanism underlying this modulation but further evidence is required to support this conclusion. The abundance of successful applications of non-invasive brain stimulation techniques for investigating neuronal oscillations in auditory processing is promising for future studies of oscillatory activity and AM sensitivity.

To summarise, non-invasive brain stimulation techniques have many effective applications relating to neural oscillations in speech and auditory processing. Indeed, these methods are uniquely positioned for non-invasively providing causal evidence for the role of oscillatory activity. tCS techniques is amenable to the goal of influencing or directly entraining oscillations due to their low stimulation intensity and their ability to modulate timing and neuronal firing rate (while not directly initiating firing like TMS). TMS shares some desirable properties with tCS and as the effects of TMS on neural oscillations are further elucidated, their application is sure to proliferate.

6.4.3 Neural Oscillations and Predictions in Speech

One key aspect underlying the current interest in oscillatory patterns of neuronal activity is the suggestion that their rhythmic temporal nature serves to optimise processing (Arnal and Giraud 2012; Morillon and Schroeder 2015; Nobre and van Ede 2018). Empirical evidence for this has come from studies where stimulus predictability was modulated while neuronal oscillations were measured. It has also been shown behaviourally that temporal predictability of low-level visual or auditory stimuli can enhance detection (Rohenkohl et al. 2012; Lawrance et al. 2014). Rhythmic sensory stimuli can be used to actively entrain the frequency and phase of neuronal oscillations and this has been shown to improve subsequent perception (Calderone et al. 2014; Hickok et al. 2015). Musical rhythms have been used to improve subsequent behavioural performance and this increase was associated with modulations of the EEG signal (Falk et al. 2017). At a higher level, it has also been suggested that this predictability improves language processing and evidence has been presented for pre-activation of semantic and phonological representations (DeLong et al. 2005). However, this is contentious as a recent multi-site study failed to replicate this original finding (Nieuwland et al. 2018).

Rhythmic sensory or electrical stimulation facilitates processing of subsequent low-level stimuli however, the exact mechanism subserving these behavioural improvements is not known. Predictive coding accounts have been proposed to explain this facilitation and there is mounting support for predictive coding in auditory cortex (Bendixen et al. 2012; Aitchison and Lengyel 2017; Heilbron and Chait 2017). There are multiple models of predictive coding (Spratling 2017) but their key tenet is that the brain continuously predicts sensory expectations and cross-references them with sensory input (Clark 2013). A prediction error occurs when there is a misalignment between observed sensory input and the predicted response from an internal generative model. These prediction errors are surprising in the formal sense, the most informative parts of sensory input, so physiologically-expensive spiking energy can be used more efficiently if only these prediction errors are encoded. The internal model is amended in response to prediction errors and many predictive coding accounts posit Bayesian model updating as a statistical formulation of this process. This is proposed to have an hierarchical structure with separate neural populations encoding top-down feedforward predictions and bottom-up feedback of prediction errors (Wacongne et al. 2011). These distinct processes are also thought to manifest in different neural oscillatory bands, with predictions linked to low gamma and prediction errors linked to beta. Neuroanatomical models have been developed that account for extracellular response profiles similar to those measured empirically (Rao and Ballard 1999). Key to one prominent instantiation of predictive coding and Bayesian inferential modelling is the simple principle that the brain optimises processing by minimisation of prediction error (Friston 2010). It has recently been proposed that sensory cortical regions may encode the precision of prediction errors, and this optimises the model sensitivity to sensory input (by manipulating its Bayesian prior distribution) (Hesselmann et al. 2010; Friston 2018). This general theory of brain function parsimoniously explains information processing across action, perception and learning across traditionally disparate higher-level psychological processes (Friston 2003; Clark 2013; Haegens and Zion Golumbic 2018). Despite the growing theoretical, neuroimaging and computational modelling support, key claims of predictive coding accounts are difficult or impossible to test without more direct cellular work and more detailed accounts of how it is neurophysiologically instantiated (Kogo and Trengove 2015). Care must also be taken to distinguish between predictive coding and attentional processes that may be reflected in biases (e.g. to specific frequency bands) but do not directly reflect prediction (Hsu et al. 2014). These aspects are difficult to dissociate and are often conflated leading to unclear outcomes (Heilbron and Chait 2017). Emphasis should be placed on careful experimental design to ensure specificity of measurement to minimise possible confounding attentional factors. Non-invasive measures such as indices of surprise (e.g. mismatch negativity) and entrainment metrics reflect the output of multiple psychological processes and many neural populations. It is important to consider the possible mismatch between these coding accounts and the specificity of measurements non-invasive imaging techniques allow.

Measures of low-level auditory processing have been fundamental in developing models of predictive coding. The success of these applications to low-level audition is encouraging for future investigations of AM processing. The rhythmic and predictable nature of AM of the speech envelope is likely to exploited by the brain to optimise processing, as has been shown for many low level aspects in vision and audition. Further, the strong evidence for the involvement of neural oscillations in the processing of low frequency envelope information suggests that it may be fruitful to apply predictive coding concepts and dynamic causal modelling in order to further understand AM processing. There is a paucity of research applying non-invasive brain stimulation methods to specifically test accounts of predictive coding. This is likely underpinned by a lack of mechanistic understanding and a lack of specificity of non-invasive brain stimulation effects. Predictive coding is thought to be instantiated between different cortical layers, whereas brain stimulation techniques are most effective for stimulating broad brain regions or nodes at the macroscopic systems level. For example, the argument that distinct subpopulations of neurons represent predictions and prediction errors may be testable using non-invasive brain stimulation in future, but new methods with increased stimulation specificity are required. Though testing these accounts may not be possible with non-invasive brain stimulation, previous findings in the literature can be reinterpreted using predictive coding concepts. For example, studies demonstrating the state-dependent TMS have used selective behavioural adaptation combined with TMS pulses to influence behaviour (Cattaneo et al. 2008: e.g). These adaptation effects are commonly thought to be due to fatigue of neural systems tuned to detect the stimulus, however it may actually be a reflection of distributional learning (Kleinschmidt and Jaeger 2016).

To summarise, behavioural and EEG studies have shown that rhythmicity in sensory or electrical stimulation enables the brain to process subsequent stimuli more accurately. A prominent explanation for this phenomenon is that predictive coding plays a role in the facilitation by increasing the efficiency of processing. Minimisation of prediction error is proposed to parsimoniously account for many brain functions underlying action, perception and learning. The evidence for predictive coding in auditory cortex is unclear, however many low-level auditory processes have been used to investigate this. It is promising for future studies of AM processing that low-level audition is commonly used as a testbed for developing new models of predictive coding.

6.4.4 Speech Production and Perception

The neurobiological networks underlying speech perception and production share many nodes and speech production mechanisms are intrinsically linked to speech perception. In the general approach to this thesis, auditory processing has been primarily conceptualised as a spectrotemporal decomposition and analysis of acoustic input, subserved by auditory cortical regions. This may be sufficient for low-level controlled stimuli, but explanations of higher order processing must also address the role of the motor system. As an illustrative example, it has recently been found that speakers enhance their production of low frequency AM when speaking in noisy conditions (Bosker and Cooke 2018). This concurs with arguments for the importance of speech envelope AM for intelligibility but from the perspective of the speaker enhancing them in order to be understood. Here it serves to emphasise how intrinsically connected these low-level components are to the articulatory motor apparatus that produces them. Indeed, the syllabic rate (∼4 Hz) has been proposed to be the natural oscillating rate of the jaw (Ohala 1975). Although it is difficult to determine the direction of cause and effect with this observation, the articulatory apparatus clearly poses some limits on this rate due to physical constraints. Behavioural approaches like this have contributed a remarkable amount to our understanding of speech processing. Speakers continuously adapt to their acoustic environment, including their own vocalisations, and atypicalities within this system have been linked to disorders with speech processing deficits, such as autism (Lee 1950; Jones and Striemer 2007; Lin et al. 2015; Luo et al. 2018). Systematically manipulating the adverse conditions and measuring how speech articulation changes can be an informative behavioural technique for investigating how speech production is altered when intelligibility is low (e.g. Hazan and Baker 2011).

The neural instantiation of the speech-motor network is fairly well characterised (Hickok 2012). However, it is controversial how much activation of the motor system underlies the perception of speech. An early integrative theory posited that listeners understand speech by detecting their intended articulatory gestures, rather than directly processing acoustic input (Liberman et al. 1967; Liberman and Mattingly 1985). This was developed to explain co-articulation; where a phoneme has different acoustic properties depending on the phonemes that precede and follow it. Co-articulation is a common part of natural speech that does not hinder intelligibility, so it was thought that invariant motor representations used to produce phonemes were activated in the listener. While a strong version of this theory has been largely rejected (Möttönen and Watkins 2012; Hickok et al. 2011), it was the first to implicate motor processes in speech perception and this continues to be near the forefront of speech perception theory (e.g. Wilson and Knoblich 2005; Pickering and Garrod 2013). A dual-stream model has been proposed to account for speech perception and production and the nodes that overlap both functions. In the left hemisphere, speech perception is proposed to be driven primarily via a dorsal stream and speech production dominated by activity in a ventral stream (Hickok and Poeppel 2007). An asymmetrical coupling between speech perception and production neural systems was measured using concurrent fMRI-EEG and it was proposed that this may underlie the hemispheric lateralisation of speech (Giraud et al. 2007). This information has also been incorporated into quantitative computational models of the motor articulatory system that implicate overlapping speech perception and production networks (Guenther and Vladusich 2012). A particularly compelling account has been proposed for the role of the the speech production system in prediction of speech (Martin et al. 2018). This event-related potential EEG study found that prediction of subsequent words was only disrupted when the speech production system was taxed. This aligns with the described role of the motor system in predicting the state of the vocal tract and of sensory representations from one's own speech (Hickok et al. 2011). This results in decreased activations in auditory areas when speech is self-generated, compared to when a recording of the same acoustic stimulus is played (Aliu et al. 2008). Although this novel extension to speech which is not self-generated requires further investigation, it is promising initial evidence for another possible role of motor areas in speech perception.

Non-invasive brain stimulation techniques have been crucial in investigating the basis for a causal role of speech motor regions for speech perception (Möttönen and Watkins 2012; Murakami et al. 2013; Schomers and Pulvermüller 2016). Measuring motor evoked potentials (MEPs) from the lips has been shown to be a useful technique for measuring excitability of speech-related motor areas (Watkins et al. 2003). It was shown that lip MEPs were reduced in conditions where participants heard speech or viewed speech-related lip movements. Further, these excitability changes have been found to be reflected in the BOLD response (Pulvermüller et al. 2006), although this is not always the case, possibly due to interindividual differences in the location of speech-related motor areas (Möttönen and Watkins 2012). 'Virtual lesion' TMS protocols have also been applied to investigate the role of the motor system in speech perception (Möttönen and Watkins 2009). Meister et al. (2007) showed that rTMS to premotor cortex impaired phoneme discrimination sensitivity compared to a colour discrimination control task. They also found a larger decrease in phoneme discrimination accuracy with TMS to the premotor cortex than with TMS to the superior temporal gyrus. TMS has provided some evidence for the causal role of the motor cortex in speech perception, however this is controversial. These claims have been criticised as motor responses (button presses) were required, which is dissimilar to natural speech processing and may confound the results (Schomers and Pulvermüller 2016). It is also possible that TMS influences secondary, decision-making processes as opposed to the direct perception of speech. A particularly robust finding is that activation in articulatory motor areas is facilitated under adverse listening conditions (Wilson and Knoblich 2005; Adank and Devlin 2010; Murakami et al. 2011; Adank et al. 2013; Nuttall et al. 2016, 2017). The current evidence supports the idea that the motor cortex has a clear modulatory effect on speech processing, particularly under adverse listening conditions, but may not be essential for understanding clear speech(Sato et al. 2009; Hickok et al. 2011).

Investigations into the role of articulatory motor areas in speech perception can help to inform future studies of AM processing. Although there is evidence for a causal role of the speech motor system in some speech-related processes, such as phoneme discrimination (D'Ausilio et al. 2009), it is not clear if an AM signal that approximates speech would involve changes in motor system activation. Investigating the effects found in this thesis further with more ecologically valid stimuli would be an interesting line of future enquiry. For example, changing the AM shape from sinusoidal to pulsatile bursts that more closely approximate AM in speech (Prendergast et al. 2010). Further, these modulations could be parametrically varied from sinusoidal AM to a full speech envelope and the full response profile characterised with varying AM depths to control how challenging the listening conditions were. It would be interesting to determine if there is a key parameter or AM rate that makes a controlled signal sufficiently 'speech-like' as to involve motor processes. TMS could be delivered to jaw motor areas while EMG is measured from the digastric muscles (Sowman et al. 2009) to see if similar disruptions found for speech processing and lip M1 are also found for AM processing and jaw M1. Exploring this with different AM rates would also be interesting as a somatotopic effect has been shown where lip M1 MEPs were facilitated more when the noise-distortion was due to lip restriction, rather than tongue restriction (Nuttall et al. 2016). Following this, it may be predicted that the band-pass TMTF shape with peak sensitivity at ∼2 – 4 Hz (Edwards and Chang 2013) would be measured from the excitability of jaw motor areas as AM rates outside this range are less relevant to the targeted area. It is, however, also possible that modification to the AM shape alone would not result in changes in motor excitability, suggesting that more complex speech components (e.g. spectrotemporal modulations) may be required for motor cortex involvement. The contribution of the motor cortex to speech perception increases under more challenging listening conditions (e.g. Adank and Devlin 2010), so there may be a non-linear response profile where motor activation is not apparent for sinusoidal AM noise but increases with speech envelope-AM noise, then decreases again as more aspects are added and intelligibility improves. This would need to be carefully controlled as it may be the quantity of noise distortion, rather than the quality (speech relevant or non-speech relevant distortions) that predominantly mediates the involvement of motor regions and different response profiles have been shown for participants with different hearing ability, even within the limits of normal hearing (Nuttall et al. 2017). An alternative approach would be to start with the full speech signal and then progressively remove low-level acoustic aspects, to quantify how activation changes and which parts are specific to speech.

To summarise, speech production and speech perception are highly interlinked. Cortical processing of low-level speech aspects primarily involves auditory cortex, however there is a much wider network involved in speech processing, including motor regions. Future studies that aim to expand the TMS effects on AM processing to speech processing generally must account for this. There is some evidence for a causal role of articulatory motor areas to speech perception but this is controversial. Non-invasive brain stimulation techniques have been valuable for exploring this further. These studies have shown that the articulatory motor regions are most active under challenging listening conditions, suggesting a modulatory role of M1 on speech processing. Applying these techniques to investigate AM processing might provide some novel insights and clear testable predictions can be made from this literature.

6.5 Follow-up Work and Future Directions

The work conducted in this thesis was fairly exploratory in nature as before this, it was not known whether it was possible to affect AM sensitivity with TMS. While this may have led to some unexpected findings, caveats and difficult interpretations, this initial work will help to inform future investigations aiming to elucidate the mechanisms underlying AM processing. Indeed, the studies completed as part of this project have already informed subsequent work within the research group.

One current arm of research is aiming to improve the adaptive staircasing and function fitting through comparison of different psychophysical designs using naïve participants. The task designs include the two main methods used here and additional staircase types including simple up/down and Bayesian algorithms (Levitt 1971; Watson and Pelli 1983; Watson 2017). This work will also explore the implementation of updated maximum likelihood and the effects of adding lapse rate and guess rate terms to the psychometric function (Shen and Richards 2013; Shen et al. 2015; Wichmann and Hill 2001).

Another line of research informed by this work is investigating the efficacy of TMS to dissociate peripheral and cortical deficits in populations with normal hearing, sensorineural hearing loss, noise-exposure and across different age brackets. As TMS directly affects cortical regions (and is now known to be effective at modulating AM processing), it is being applied to dissociate deficits due to central and peripheral factors. If TMS proves effective it could be used alone or as a complement to other newly developed methods that aim to elucidate the mechanisms underlying hidden hearing loss (Dewey et al. 2018). This could be further extended by comparing the functional connectivity profiles of people with normal hearing and those with different types of hearing impairment or at different stages of hearing loss.

Finally, the behavioural data collected as part of these studies is also being used to inform a larger scale cohort study investigating the relationship between AM detection and discrimination ability and functional or structural connectivity measured during resting state. It is known that the connectivity of auditory areas is asymmetric (Mišić et al. 2018), however, this has not been linked to AM detection and discrimination performance. It would be particularly interesting to explore whether this connectivity plays a role in AM sensitivity in normal hearing adults and if so, if this is reflected in the connectivity profile. For example, is finer AM sensitivity associated with a greater degree of interhemispheric connectivity, and is this pattern the same for sensitivity to 4 and 40 Hz AM?

6.6 Conclusions

This research was designed to address the following three overarching research questions:

- Firstly, is fMRI-guided TMS an effective method for modulating AM processing?
- Secondly, are different TMS protocols more or less effective at modulating AM processing?
- Finally, is fMRI-guided TMS an effective method for further understanding the functional asymmetry of speech processing?

It was shown that TMS can be an effective method for modulating AM sensitivity, an integral part of speech processing. It was also shown that different TMS protocols vary in effectiveness and possible explanations for this were discussed. Finally, it was shown that TMS can be an effective tool for further understanding the functional asymmetry of speech processing, however some caveats must be accounted for. This work has opened many new avenues for future research to explore and made key recommendations for similar investigations. Recent developments in psychophysical, fMRI and TMS methods could all be incorporated to improve future studies. Overall, the outlook is promising for research investigating AM processing using fMRI-guided TMS. To conclude, there is clearly still much to learn about how AM is processed cortically, however, this thesis has shown that fMRI-guided TMS can be an effective method by which to do so.

References

Abdul-Kareem IA, Sluming V (2008) Heschl gyrus and its included primary auditory cortex: Structural MRI studies in healthy and diseased subjects. *Journal of Magnetic Resonance Imaging* 28:287–299.

Abrams DA, Nicol T, Zecker S, Kraus N (2008) Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *Journal of Neuroscience* 28:3958–3965.

Abrams DA, Nicol T, Zecker S, Kraus N (2009) Abnormal Cortical Processing of the Syllable Rate of Speech in Poor Readers. *Journal of Neuroscience* 29:7686–7693.

Adank P, Devlin JT (2010) On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage* 49:1124–1132.

Adank P, Nuttall HE, Kennedy-Higgins D (2017) Transcranial magnetic stimulation and motor evoked potentials in speech perception research. *Language, Cognition and Neuroscience* 32:900–909.

Adank P, Rueschemeyer SA, Bekkering H (2013) The role of accent imitation in sensorimotor integration during processing of intelligible speech. *Frontiers in Human Neuroscience* 7.

Ainsworth M, Lee S, Cunningham MO, Traub RD, Kopell NJ, Whittington MA (2012) Rates and Rhythms: A Synergistic View of Frequency and Temporal Coding in Neuronal Networks. *Neuron* 75:572–583.

Aitchison L, Lengyel M (2017) With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology* 46:219–227.

Albouy P, Weiss A, Baillet S, Zatorre RJ (2017) Selective Entrainment of Theta Oscillations in the Dorsal Stream Causally Enhances Auditory Working Memory Performance. *Neuron* 94:193–206.

Aliu SO, Houde JF, Nagarajan SS (2008) Motor-induced Suppression of the Auditory Cortex. *Journal of Cognitive Neuroscience* 21:791–802.

Amitay S, Irwin A, Hawkey DJC, Cowan JA, Moore DR (2006) A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naive listeners. *The Journal of the Acoustical Society of America* 119:1616–1625.

Andoh J, Ferreira M, Leppert IR, Matsushita R, Pike B, Zatorre RJ (2017) How restful is it with all that noise? Comparison of Interleaved silent steady state (ISSS) and conventional imaging in resting-state fMRI. *NeuroImage* 147:726–735.

Andoh J, Matsushita R, Zatorre RJ (2015) Asymmetric Interhemispheric Transfer in the Auditory Network: Evidence from TMS, Resting-State fMRI, and Diffusion Imaging. *Journal of Neuroscience* 35:14602–14611.

Andoh J, Zatorre RJ (2011) Interhemispheric Connectivity Influences the Degree of Modulation of TMS-Induced Effects during Auditory Processing. *Frontiers in Psychology* 2.

Andoh J, Zatorre RJ (2013) Mapping interhemispheric connectivity using functional MRI after transcranial magnetic stimulation on the human auditory cortex. *NeuroImage* 79:162–171.

Antal A, Paulus W (2013) Transcranial alternating current stimulation (tACS). *Frontiers in Human Neuroscience* 7.

Arnal LH, Giraud AL (2012) Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences* 16:390–398.

Baayen RH, Davidson DJ, Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59:390–412.

Bacon SP, Grantham DW (1989) Modulation masking: Effects of modulation frequency, depth, and phase. *The Journal of the Acoustical Society of America* 85:2575–2580.

Bak JH, Pillow JW (2018) Adaptive stimulus selection for multi-alternative psychometric functions with lapses. *bioRxiv* p. 260976.

Baltus A, Herrmann CS (2015) Auditory temporal resolution is linked to resonance frequency of the auditory cortex. *International Journal of Psychophysiology* 98:1–7.

Baltus A, Wagner S, Wolters CH, Herrmann CS (2018) Optimized auditory transcranial alternating current stimulation improves individual auditory temporal resolution. *Brain Stimulation* 11:118–124.

Barker AT, Jalinous R, Freeston IL (1985) Non-Invasive Magnetic Stimulation of Human Motor Cortex. *The Lancet* 325:1106–1107.

Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language* 68:255–278.

Barton B, Venezia JH, Saberi K, Hickok G, Brewer AA (2012) Orthogonal acoustic dimensions define auditory field maps in human cortex. *Proceedings of the National Academy of Sciences of the United States of America* 109:20738–20743.

Bates D, Machler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67.

Bausenhart KM, Dyjas O, Ulrich R (2015) Effects of stimulus order on discrimination sensitivity for short and long durations. *Attention, Perception, & Psychophysics* 77:1033–1043.

Behler O, Uppenkamp S (2016) The representation of level and loudness in the central auditory system for unilateral stimulation. *NeuroImage* 139:176–188.

Bendixen A, SanMiguel I, Schröger E (2012) Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology* 83:120–131.

Bergmann TO, Karabanov A, Hartwigsen G, Thielscher A, Siebner HR (2016) Combining non-invasive transcranial brain stimulation with neuroimaging and electrophysiology: Current approaches and future perspectives. *NeuroImage* 140:4–19.

Berker EA, Berker AH, Smith A (1986) Translation of Broca's 1865 Report: Localization of Speech in the Third Left Frontal Convolution. *Archives of Neurology* 43:1065–1072.

Bestmann S, Baudewig J, Siebner HR, Rothwell JC, Frahm J (2003) Subthreshold high-frequency TMS of human primary motor cortex modulates interconnected frontal motor areas as detected by interleaved fMRI-TMS. *NeuroImage* 20:1685–1696.

Bestmann S, Baudewig J, Siebner HR, Rothwell JC, Frahm J (2004) Functional MRI of the immediate impact of transcranial magnetic stimulation on cortical and subcortical motor circuits. *European Journal of Neuroscience* 19:1950–1962.

Bestmann S, Baudewig J, Siebner HR, Rothwell JC, Frahm J (2005) BOLD MRI responses to repetitive TMS over human dorsal premotor cortex. *NeuroImage* 28:22–29.

Bestmann S, de Berker AO, Bonauito J (2015) Understanding the behavioural consequences of noninvasive brain stimulation. *Trends in Cognitive Sciences* 19:13–20.

Bishop GH (1932) Cyclic changes in excitability of the optic pathway of the rabbit. *American Journal of Physiology-Legacy Content* 103:213–224.

Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience* 8:389–395.

Bosker HR, Cooke M (2018) Talkers produce more pronounced amplitude modulations when speaking in noise. *The Journal of the Acoustical Society of America* 143:EL121–EL126.

Bosker HR, Ghitza O (2018) Entrained theta oscillations guide perception of subsequent speech: Behavioural evidence from rate normalisation. *Language, Cognition and Neuroscience* 33:1–13.

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA (2009) Coding of Repetitive Transients by Auditory Cortex on Heschl's Gyrus. *Journal of Neurophysiology* 102:2358–2374.

Bueti D, van Dongen EV, Walsh V (2008) The Role of Superior Temporal Cortex in Auditory Timing. *PLOS ONE* 3:e2481.

Calderone DJ, Lakatos P, Butler PD, Castellanos FX (2014) Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences* 18:300–309.

Cammoun L, Thiran JP, Griffa A, Meuli R, Hagmann P, Clarke S (2015) Intrahemispheric cortico-cortical connections of the human auditory cortex. *Brain Structure and Function* 220:3537–3553.

Carbajal GV, Malmierca MS (2018) The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection. *Trends in Hearing* 22:1–33.

Catani M, Allin MPG, Husain M, Pugliese L, Mesulam MM, Murray RM, Jones DK (2007) Symmetries in human brain language pathways correlate with verbal recall. *Proceedings of the National Academy of Sciences* 104:17163–17168.

Catani M, Jones DK, Ffytche DH (2005) Perisylvian language networks of the human brain. *Annals of Neurology* 57:8–16.

Cattaneo Z, Rota F, Vecchi T, Silvanto J (2008) Using state-dependency of transcranial magnetic stimulation (TMS) to investigate letter selectivity in the left posterior parietal cortex: A comparison of TMS-priming and TMS-adaptation paradigms. *European Journal of Neuroscience* 28:1924–1929.

Cattaneo Z, Silvanto J (2008) Time course of the state-dependent effect of transcranial magnetic stimulation in the TMS-adaptation paradigm. *Neuroscience Letters* 443:82–85.

Chait M, Greenberg S, Arai T, Simon JZ, Poeppel D (2015) Multi-time resolution analysis of speech: Evidence from psychophysics. *Frontiers in Neuroscience* 9.

Chi T, Gao Y, Guyton MC, Ru P, Shamma S (1999) Spectro-temporal modulation transfer functions and speech intelligibility. *The Journal of the Acoustical Society of America* 106:2719–2732.

Chiarello C, Vazquez D, Felton A, McDowell A (2016) Structural asymmetry of the human cerebral cortex: Regional and between-subject variability of surface area, cortical thickness, and local gyrification. *Neuropsychologia* 93:365–379.

Chung SW, Hill AT, Rogasch NC, Hoy KE, Fitzgerald PB (2016) Use of theta-burst stimulation in changing excitability of motor cortex: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews* 63:43–64.

Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36:181–204.

Cutini S, Szűcs D, Mead N, Huss M, Goswami U (2016) Atypical right hemisphere response to slow temporal modulations in children with developmental dyslexia. *NeuroImage* 143:40–49.

Dau T, Kollmeier B, Kohlrausch A (1997a) Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *The Journal of the Acoustical Society of America* 102:2892–2905.

Dau T, Kollmeier B, Kohlrausch A (1997b) Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *The Journal of the Acoustical Society of America* 102:2906–2919.

D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L (2009) The Motor Somatotopy of Speech Perception. *Current Biology* 19:381–385.

Davis MH, Johnsrude IS (2003) Hierarchical Processing in Spoken Language Comprehension. *Journal of Neuroscience* 23:3423–3431.

DeLong KA, Urbach TP, Kutas M (2005) Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* 8:1117–1121.

Dewey RS, Hall DA, Guest H, Prendergast G, Plack CJ, Francis ST (2018) The Physiological Bases of Hidden Noise-Induced Hearing Loss: Protocol for a Functional Neuroimaging Study. *JMIR Research Protocols* 7:e79.

Dick FK, Lehet MI, Callaghan MF, Keller TA, Sereno MI, Holt LL (2017) Extensive Tonotopic Mapping across Auditory Cortex Is Recapitulated by Spectrally Directed Attention and Systematically Related to Cortical Myeloarchitecture. *The Journal of Neuroscience* 37:12187–12201.

Ding N, Melloni L, Zhang H, Tian X, Poeppel D (2016) Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience* 19:158–164.

Ding N, Patel AD, Chen L, Butler H, Luo C, Poeppel D (2017) Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews* 81:181–187.

Dixon P (2008) Models of accuracy in repeated-measures designs. *Journal of Memory and Language* 59:447–456.

Doelling KB, Arnal LH, Ghitza O, Poeppel D (2014) Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage* 85:761–768.

Doll RJ, Veltink PH, Buitenweg JR (2015) Observation of time-dependent psychophysical functions and accounting for threshold drifts. *Attention, Perception, & Psychophysics* 77:1440–1447.

Dorman MF, Loizou PC, Rainey D (1997) Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *The Journal of the Acoustical Society of America* 102:2403–2411.

Drullman R, Festen JM, Plomp R (1994a) Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America* 95:2670–2680.

Drullman R, Festen JM, Plomp R (1994b) Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America* 95:1053–1064.

Dyjas O, Ulrich R (2014) Effects of Stimulus Order on Discrimination Processes in Comparative and Equality Judgements: Data and Models. *Quarterly Journal of Experimental Psychology* 67:1121–1150.

Edwards E, Chang EF (2013) Syllabic (∼2–5 Hz) and fluctuation (∼1–10 Hz) ranges in speech and auditory processing. *Hearing Research* 305:113–134.

Eldaief MC, Halko MA, Buckner RL, Pascual-Leone A (2011) Transcranial magnetic stimulation modulates the brain's intrinsic activity in a frequency-dependent manner. *Proceedings of the National Academy of Sciences* 108:21229–21234.

Elliott MR, Bowtell RW, Morris PG (1999) The effect of scanner sound in visual, motor, and auditory functional MRI. *Magnetic Resonance in Medicine* 41:1230–1235.

Elliott TM, Theunissen FE (2009) The Modulation Transfer Function for Speech Intelligibility. *PLOS Computational Biology* 5:e1000302.

Ewert SD, Dau T (2000) Characterizing frequency selectivity for envelope fluctuations. *The Journal of the Acoustical Society of America* 108:1181–1196.

Ewert SD, Dau T (2004) External and internal limitations in amplitude-modulation processing. *The Journal of the Acoustical Society of America* 116:478–490.

Falk S, Lanzilotti C, Schön D (2017) Tuning Neural Phase Entrainment to Speech. *Journal of Cognitive Neuroscience* 29:1378–1389.

Farzan F, Vernet M, Shafi MMD, Rotenberg A, Daskalakis ZJ, Pascual-Leone A (2016) Characterizing and Modulating Brain Circuitry through Transcranial Magnetic Stimulation Combined with Electroencephalography. *Frontiers in Neural Circuits* 10.

Fitzgibbons PJ, Wightman FL (1982) Gap detection in normal and hearing-impaired listeners. *The Journal of the Acoustical Society of America* 72:761–765.

Fleischer H (1980) Subjektive Größe von Unterschieden im Amplituden-Modulationsgrad von Sinustönen. *Acustica* 46:31–38.

Fox MD, Buckner RL, Liu H, Chakravarty MM, Lozano AM, Pascual-Leone A (2014) Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. *Proceedings of the National Academy of Sciences of the United States of America* 111:E4367–E4375.

Frank SL, Yang J (2018) Lexical representation explains cortical entrainment during speech comprehension. *PLOS ONE* 13:e0197304.

Frisby JP (1979) *Seeing : Illusion, Brain, and Mind* Oxford University Press, Oxford, UK.

Frisina RD (2001) Subcortical neural coding mechanisms for auditory temporal processing. *Hearing Research* 158:1–27.

Friston K (2003) Learning and inference in the brain. *Neural Networks* 16:1325–1352.

Friston K (2010) The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience* 11:127–138.

Friston K (2018) Does predictive coding have a future? *Nature Neuroscience* 21:1019–1021.

Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences* 78:2643–2647.

Galuske RAW, Schlote W, Bratzke H, Singer W (2000) Interhemispheric Asymmetries of the Modular Structure in Human Temporal Cortex. *Science* 289:1946–1949.

Gao L, Kostlan K, Wang Y, Wang X (2016) Distinct Subthreshold Mechanisms Underlying Rate-Coding Principles in Primate Auditory Cortex. *Neuron* 91:905–919.

García-Pérez MA, Alcalá-Quintana R (2013) Shifts of the psychometric function: Distinguishing bias from perceptual effects , Shifts of the psychometric function: Distinguishing bias from perceptual effects. *Quarterly Journal of Experimental Psychology* 66:319–337.

Geeter ND, Dupré L, Crevecoeur G (2016) Modeling transcranial magnetic stimulation from the induced electric fields to the membrane potentials along tractography-based white matter fiber tracts. *Journal of Neural Engineering* 13:026028.

Geschwind N, Levitsky W (1968) Human Brain: Left-Right Asymmetries in Temporal Speech Region. *Science* 161:186–187.

Ghitza O (2011) Linking Speech Perception and Neurophysiology: Speech Decoding Guided by Cascaded Oscillators Locked to the Input Rhythm. *Frontiers in Psychology* 2.

Ghitza O (2017) Acoustic-driven delta rhythms as prosodic markers. *Language, Cognition and Neuroscience* 32:545–561.

Ghitza O, Giraud AL, Poeppel D (2013) Neuronal oscillations and speech perception: Critical-band temporal envelopes are the essence. *Frontiers in Human Neuroscience* 6.

Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RSJ, Laufs H (2007) Endogenous Cortical Rhythms Determine Cerebral Specialization for Speech Perception and Production. *Neuron* 56:1127–1134.

Giraud AL, Lorenzi C, Ashburner J, Wable J, Johnsrude I, Frackowiak R, Kleinschmidt A (2000) Representation of the temporal envelope of sounds in the human brain. *Journal of neurophysiology* 84:1588–1598.

Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature neuroscience* 15:511–517.

Gold T, Pumphrey RJ (1948) Hearing. I. The Cochlea as a Frequency Analyzer. *Proceedings of the Royal Society of London B: Biological Sciences* 135:462–491.

Goodwin BD, Butson CR (2015) Subject-Specific Multiscale Modeling to Investigate Effects of Transcranial Magnetic Stimulation. *Neuromodulation: Technology at the Neural Interface* 18:694–704.

Goswami U, Leong V (2013) Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology* 4:67–92.

Grantham DW, Bacon SP (1988) Detection of increments and decrements in modulation depth of SAM noise. *The Journal of the Acoustical Society of America* 84:S140–S140.

Gratton C, Lee TG, Nomura EM, D'Esposito M (2014) Perfusion MRI Indexes Variability in the Functional Brain Effects of Theta-Burst Transcranial Magnetic Stimulation. *PLoS ONE* 9.

Green DM (1990) Stimulus selection in adaptive psychophysical procedures. *The Journal of the Acoustical Society of America* 87:2662–2674.

Greenberg S (2004) What are the essential cues for understanding spoken language? In *IEICE Transactions on Information and Systems*, Vol. 109, pp. 2382–2382.

Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. *Trends in Neurosciences* 25:348–353.

Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences* 10:14–23.

Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S (2013) Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLOS Biology* 11:e1001752.

Grothe B, Pecka M, McAlpine D (2010) Mechanisms of Sound Localization in Mammals. *Physiological Reviews* 90:983–1012.

Guenther FH, Vladusich T (2012) A neural theory of speech acquisition and production. *Journal of Neurolinguistics* 25:408–422.

Guest H, Munro KJ, Prendergast G, Millman RE, Plack CJ (2018) Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hearing Research* 364:142–151.

Gutschalk A, Micheyl C, Oxenham AJ (2008) The pulse-train auditory aftereffect and the perception of rapid amplitude modulations. *The Journal of the Acoustical Society of America* 123:935–945.

Haegens S, Zion Golumbic E (2018) Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews* 86:150–165.

Halekoh U, Hojsgaard S (2014) A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models: The R Package pbkrtest. *Journal of Statistical Software* 59.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999) ”Sparse” temporal sampling in auditory fMRI. *Human Brain Mapping* 7:213–223.

Hallam GP, Whitney C, Hymers M, Gouws AD, Jefferies E (2016) Charting the effects of TMS with fMRI: Modulation of cortical recruitment within the distributed network supporting semantic control. *Neuropsychologia* 93:40–52.

Hamilton LS, Huth AG (2018) The revolution will not be controlled: Natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience* 0:1–13.

Han JH, Dimitrijevic A (2015) Acoustic change responses to amplitude modulation: A method to quantify cortical temporal processing and hemispheric asymmetry. *Frontiers in Neuroscience* 9.

Harms MP, Guinan JJ, Sigalovsky IS, Melcher JR (2005) Short-Term Sound Temporal Envelope Characteristics Determine Multisecond Time Patterns of Activity in Human Auditory Cortex as Shown by fMRI. *Journal of Neurophysiology* 93:210–222.

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 (2003) Amplitude and Frequency-modulated Stimuli Activate Common Regions of Human Auditory Cortex. *Cerebral Cortex* 13:773–781.

Hartwigsen G (2015) The neurophysiology of language: Insights from non-invasive brain stimulation in the healthy human brain. *Brain and Language* 148:81–94.

Hazan V, Baker R (2011) Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America* 130:2139–2152.

Heilbron M, Chait M (2017) Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience* .

Heimrath K, Kuehne M, Heinze HJ, Zaehle T (2014) Transcranial direct current stimulation (tDCS) traces the predominance of the left auditory cortex for processing of rapidly changing acoustic information. *Neuroscience* 261:68–73.

Heimrath K, Fiene M, Rufener KS, Zaehle T (2016) Modulating Human Auditory Processing by Transcranial Electrical Stimulation. *Frontiers in Cellular Neuroscience* 10.

Herdener M, Esposito F, Scheffler K, Schneider P, Logothetis NK, Uludag K, Kayser C (2013) Spatial representations of temporal and spectral sound cues in human auditory cortex. *Cortex* 49:2822–2833.

Herring JD, Thut G, Jensen O, Bergmann TO (2015) Attention Modulates TMS-Locked Alpha Oscillations in the Visual Cortex. *Journal of Neuroscience* 35:14435–14447.

Hesselmann G, Sadaghiani S, Friston KJ, Kleinschmidt A (2010) Predictive Coding or Evidence Accumulation? False Inference and Neuronal Fluctuations. *PLOS ONE* 5:e9926.

Hickok G (2012) Computational neuroanatomy of speech production. *Nature Reviews Neuroscience* 13:135–145.

Hickok G, Farahbod H, Saberi K (2015) The Rhythm of Perception: Entrainment to Acoustic Rhythms Induces Subsequent Perceptual Oscillation. *Psychological Science* 26:1006–1013.

Hickok G, Houde J, Rong F (2011) Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron* 69:407–422.

Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nature Reviews Neuroscience* 8:393–402.

Holcombe AO (2009) Seeing slow and seeing fast: Two limits on perception. *Trends in Cognitive Sciences* 13:216–221.

Hood JD (1972) Fundamentals of Identification of Sensori-neural Hearing Loss. *British Journal of Audiology* 6:21–26.

Houtgast T (1989) Frequency selectivity in amplitude-modulation detection. *The Journal of the Acoustical Society of America* 85:1676–1680.

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

Hsieh IH, Saberi K (2010) Detection of sinusoidal amplitude modulation in logarithmic frequency sweeps across wide regions of the spectrum. *Hearing research* 262:9–18.

Hsu YF, Hamalainen J, Waszak F (2014) Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. *Frontiers in Human Neuroscience* 8.

Huang YZ, Edwards MJ, Rounis E, Bhatia KP, Rothwell JC (2005) Theta Burst Stimulation of the Human Motor Cortex. *Neuron* 45:201–206.

Hudspeth AJ (2008) Making an Effort to Listen: Mechanical Amplification in the Ear. *Neuron* 59:530–545.

Hugdahl K (2011) Fifty years of dichotic listening research – Still going and going and.... *Brain and Cognition* 76:211–213.

Hugdahl K, Westerhausen R (2016) Speech processing asymmetry revealed by dichotic listening and functional brain imaging. *Neuropsychologia* 93:466–481.

Hullett PW, Hamilton LS, Mesgarani N, Schreiner CE, Chang EF (2016) Human Superior Temporal Gyrus Organization of Spectrotemporal Modulation Tuning Derived from Speech Stimuli. *The Journal of Neuroscience* 36:2014–2026.

Humphries C, Liebenthal E, Binder JR (2010) Tonotopic organization of human auditory cortex. *NeuroImage* 50:1202–1211.

Hutsler JJ (2003) The specialized structure of human language cortex: Pyramidal cell size asymmetries within auditory and language-associated regions of the temporal lobes. *Brain and Language* 86:226–242.

Hyafil A, Fontolan L, Kabdebon C, Gutkin B, Giraud AL (2015) Speech encoding by coupled cortical theta and gamma oscillations. *eLife* 4:e06213.

Hymers M, Prendergast G, Liu C, Schulze A, Young ML, Wastling SJ, Barker GJ, Millman RE (2015) Neural mechanisms underlying song and speech perception can be differentiated using an illusory percept. *NeuroImage* 108:225–233.

Jaeger TF (2008) Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of memory and language* 59:434–446.

Jepsen ML, Ewert SD, Dau T (2008) A computational model of human auditory signal processing and perception. *The Journal of the Acoustical Society of America* 124:422–438.

Jones JA, Striemer D (2007) Speech disruption during delayed auditory feedback with simultaneous visual feedback. *The Journal of the Acoustical Society of America* 122:EL135–EL141.

Joris PX, Schreiner CE, Rees A (2004) Neural Processing of Amplitude-Modulated Sounds. *Physiological Reviews* 84:541–577.

Kawashima T (2009) Attention modulates auditory adaptation produced by amplitude modulation. *The Journal of the Acoustical Society of America* 126:EL123–EL127.

Kay RH, Matthews DR (1972) On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones. *The Journal of Physiology* 225:657–677.

Kleinschmidt DF, Jaeger TF (2016) Re-examining selective adaptation: Fatiguing feature detectors, or distributional learning? *Psychonomic Bulletin & Review* 23:678–691.

Klomjai W, Katz R, Lackmy-Vallée A (2015) Basic principles of transcranial magnetic stimulation (TMS) and repetitive TMS (rTMS). *Annals of Physical and Rehabilitation Medicine* 58:208–213.

Kogo N, Trengove C (2015) Is predictive coding theory articulated enough to be testable? *Frontiers in Computational Neuroscience* 9.

Kohlrausch A, Fassel R, Dau T (2000) The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *The Journal of the Acoustical Society of America* 108:723–734.

Kong XZ, Mathias SR, Guadalupe T, ENIGMA Laterality Working Group, Glahn DC, Franke B, Crivello F, Tzourio-Mazoyer N, Fisher SE, Thompson PM, Francks C (2018) Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proceedings of the National Academy of Sciences of the United States of America* 115:E5154–E5163.

Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D (2017) Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron* 93:480–490.

Krause JC, Braida LD (2004) Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America* 115:362–378.

Kubanek J, Brunner P, Gunduz A, Poeppel D, Schalk G (2013) The Tracking of Speech Envelope in the Human Cortex. *PLOS ONE* 8:e53398.

Kujirai T, Caramia MD, Rothwell JC, Day BL, Thompson PD, Ferbert A, Wroe S, Asselman P, Marsden CD (1993) Corticocortical inhibition in human motor cortex. *The Journal of Physiology* 471:501–519.

Kumar UA, Ameenudin S, Sangamanatha AV (2012) Temporal and speech processing skills in normal hearing individuals exposed to occupational noise. *Noise & Health* 14:100–105.

Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R (1992) Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America* 89:5675–5679.

Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology* 94:1904–1911.

Lapid E, Ulrich R, Rammsayer T (2008) On estimating the difference limen in duration discrimination tasks: A comparison of the 2AFC and the reminder task. *Perception & Psychophysics* 70:291–305.

Lawrance ELA, Harper NS, Cooke JE, Schnupp JWH (2014) Temporal predictability enhances auditory detection. *The Journal of the Acoustical Society of America* 135:EL357–EL363.

Lee BS (1950) Some Effects of Side-Tone Delay. *The Journal of the Acoustical Society of America* 22:639–640.

Lee J, Bacon SP (1997) Amplitude modulation depth discrimination of a sinusoidal carrier: Effect of stimulus duration. *The Journal of the Acoustical Society of America* 101:3688–3693.

Lehongre K, Morillon B, Giraud AL, Ramus F (2013) Impaired auditory sampling in dyslexia: Further evidence from combined fMRI and EEG. *Frontiers in Human Neuroscience* 7.

Levitt H (1971) Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America* 49:467–477.

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of the speech code. *Psychological Review* 74:431–461.

Liberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21:1–36.

Licklider JCR, Webster JC, Hedlun JM (1950) On the Frequency Limits of Binaural Beats. *The Journal of the Acoustical Society of America* 22:468–473.

Liem F, Hurschler MA, Jäncke L, Meyer M (2014) On the planum temporale lateralization in suprasegmental speech perception: Evidence from a study investigating behavior, structure, and function. *Human Brain Mapping* 35:1779–1789.

Lin IF, Mochida T, Asada K, Ayaya S, Kumagaya Si, Kato M (2015) Atypical delayed auditory feedback effect and Lombard effect on speech production in high-functioning adults with autism spectrum disorder. *Frontiers in Human Neuroscience* 9.

Lisanby SH, Gutman D, Luber B, Schroeder C, Sackeim HA (2001) Sham TMS: Intracerebral measurement of the induced electrical field and the induction of motor-evoked potentials. *Biological Psychiatry* 49:460–463.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150–157.

Loizou PC, Dorman M, Tu Z (1999) On the number of channels needed to understand speech. *The Journal of the Acoustical Society of America* 106:2097–2103.

López-Alonso V, Cheeran B, Río-Rodríguez D, Fernández-del-Olmo M (2014) Inter-individual Variability in Response to Non-invasive Brain Stimulation Paradigms. *Brain Stimulation* 7:372–380.

Lorenz I, Müller N, Schlee W, Langguth B, Weisz N (2010) Short-Term Effects of Single Repetitive TMS Sessions on Auditory Evoked Activity in Patients With Chronic Tinnitus. *Journal of Neurophysiology* 104:1497–1505.

Lorenzi C, Micheyl C, Berthommier F (1995) Neuronal correlates of perceptual amplitude-modulation detection. *Hearing Research* 90:219–227.

Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature Neuroscience* 4:1131–1138.

Luo H, Poeppel D (2007) Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron* 54:1001–1010.

Luo H, Poeppel D (2012) Cortical Oscillations in Auditory Perception and Speech: Evidence for Two Temporal Windows in Human Auditory Cortex. *Frontiers in Psychology* 3.

Luo H, Wang Y, Poeppel D, Simon JZ (2006) Concurrent Encoding of Frequency and Amplitude Modulation in Human Auditory Cortex: MEG Evidence. *Journal of Neurophysiology* 96:2712–2723.

Luo J, Hage SR, Moss CF (2018) The Lombard Effect: From Acoustics to Neural Mechanisms. *Trends in Neurosciences* 1441.

Lyon RF, Katsiamis AG, Drakakis EM (2010) History and future of auditory filter models In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pp. 3809–3812.

Margulies DS, Petrides M (2013) Distinct Parietal and Temporal Connectivity Profiles of Ventrolateral Frontal Areas Involved in Language Production. *Journal of Neuroscience* 33:16846–16852.

Martin CD, Branzi FM, Bar M (2018) Prediction is Production: The missing link between language production and comprehension. *Scientific Reports* 8:1079.

May KA, Solomon JA (2013) Four Theorems on the Psychometric Function. *PLOS ONE* 8:e74815.

McGettigan C, Scott SK (2012) Cortical asymmetries in speech perception: What's wrong, what's right, and what's left? *Trends in cognitive sciences* 16:269–276.

McShane BB, Gal D, Gelman A, Robert C, Tackett JL (2017) Abandon Statistical Significance. *arXiv* p. 1709.07588.

Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M (2007) The Essential Role of Premotor Cortex in Speech Perception. *Current Biology* 17:1692–1696.

Meteyard L, Holmes NP (2018) TMS SMART – Scalp mapping of annoyance ratings and twitches caused by Transcranial Magnetic Stimulation. *Journal of Neuroscience Methods* 299:34–44.

Meyer L (2017) The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience* pp. 1–13.

Meyer L, Sun Y, Martin AE (2018) Entrainment in Disguise: The Exogenous and Endogenous Cortical Rhythms of Speech and Language Processing. *PsyArXiv* .

Meyer M, Liem F, Hirsiger S, Jäncke L, Hänggi J (2014) Cortical Surface Area and Cortical Thickness Demonstrate Differential Structural Asymmetry in Auditory-Related Areas of the Human Cortex. *Cerebral Cortex* 24:2541–2552.

Millman RE, Woods WP, Quinlan PT (2011) Functional asymmetries in the representation of noise-vocoded speech. *NeuroImage* 54:2364–2373.

Mišić B, Betzel RF, Griffa A, Reus D, A M, He Y, Zuo XN, Heuvel VD, P M, Hagmann P, Sporns O, Zatorre RJ (2018) Network-Based Asymmetry of the Human Auditory System. *Cerebral Cortex* 28:2655–2664.

Moerel M, De Martino F, Formisano E (2014) An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience* 8.

Moliadze V, Zhao Y, Eysel U, Funke K (2003) Effect of transcranial magnetic stimulation on single-unit activity in the cat primary visual cortex. *The Journal of Physiology* 553:665–679.

Moon IJ, Hong SH (2014) What Is Temporal Fine Structure and Why Is It Important? *Korean Journal of Audiology* 18:1–7.

Moore BCJ, Glasberg BR (2001) Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America* 110:1067–1073.

Morgan M, Melmoth D, Solomon J (2013) Linking hypotheses underlying Class A and Class B methods. *Visual Neuroscience* 30:197–206.

Morillon B, Lehongre K, Frackowiak RSJ, Ducorps A, Kleinschmidt A, Poeppel D, Giraud AL (2010) Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences* 107:18688–18693.

Morillon B, Schroeder CE (2015) Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences* 1337:26–31.

Möttönen R, Watkins KE (2009) Motor Representations of Articulators Contribute to Categorical Perception of Speech Sounds. *Journal of Neuroscience* 29:9819–9825.

Möttönen R, Watkins KE (2012) Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology* 26:1103–1118.

Müller BW, Stude P, Nebel K, Wiese H, Ladd ME, Forsting M, Jueptner M (2003) Sparse imaging of the auditory oddball task with functional MRI. *Neuroreport* 14:1597–1601.

Müller N, Lorenz I, Langguth B, Weisz N (2013) rTMS Induced Tinnitus Relief Is Related to an Increase in Auditory Cortical Alpha Activity. *PLOS ONE* 8:e55557.

Murakami T, Restle J, Ziemann U (2011) Observation-execution matching and action inhibition in human primary motor cortex during viewing of speech-related lip movements or listening to speech. *Neuropsychologia* 49:2045–2054.

Murakami T, Ugawa Y, Ziemann U (2013) Utility of TMS to understand the neurobiology of speech. *Frontiers in Psychology* 4.

Murphy E (2015) The brain dynamics of linguistic computation. *Frontiers in Psychology* 6.

Neuling T, Ruhnau P, Weisz N, Herrmann CS, Demarchi G (2017) Faith and oscillations recovered: On analyzing EEG/MEG signals during tACS. *NeuroImage* 147:960–963.

Nieuwland MS, Politzer-Ahles S, Heyselaar E, Segaert K, Darley E, Kazanina N, Von Grebmer Zu Wolfsthurn S, Bartolozzi F, Kogan V, Ito A, Mézière D, Barr DJ, Rousselet GA, Ferguson HJ, Busch-Moreno S, Fu X, Tuomainen J, Kulakova E, Husband EM, Donaldson DI, Kohút Z, Rueschemeyer SA, Huettig F (2018) Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife* 7:e33468.

Nitsche MA, Paulus W (2000) Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *The Journal of Physiology* 527:633–639.

Nobre AC, van Ede F (2018) Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience* 19:34–48.

Nourski KV (2017) Auditory processing in the human cortex: An intracranial electrophysiology perspective. *Laryngoscope Investigative Otolaryngology* 2:147–156.

Nourski KV, Brugge JF, Reale RA, Kovach CK, Oya H, Kawasaki H, Jenison RL, Howard MA (2012) Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: An intracranial electrophysiology study. *Journal of Neurophysiology* 109:1283–1295.

Noury N, Hipp JF, Siegel M (2016) Physiological processes non-linearly affect electrophysiological recordings during transcranial electric stimulation. *NeuroImage* 140:99–109.

Noury N, Siegel M (2018) Analyzing EEG and MEG signals recorded during tES, a reply. *NeuroImage* 167:53–61.

Nuttall HE, Kennedy-Higgins D, Devlin JT, Adank P (2017) The role of hearing ability and speech distortion in the facilitation of articulatory motor cortex. *Neuropsychologia* 94:13–22.

Nuttall HE, Kennedy-Higgins D, Hogan J, Devlin JT, Adank P (2016) The effect of speech distortion on the excitability of articulatory motor cortex. *NeuroImage* 128:218–226.

Oberman L, Edwards D, Eldaief M, Pascual-Leone A (2011) Safety of Theta Burst Transcranial Magnetic Stimulation: A Systematic Review of the Literature. *Journal of Clinical Neurophysiology* 28:67.

Obleser J, Herrmann B, Henry MJ (2012) Neural Oscillations in Speech: Don't be Enslaved by the Envelope. *Frontiers in Human Neuroscience* 6.

Obleser J, Zimmermann J, Van Meter J, Rauschecker JP (2007) Multiple Stages of Auditory Speech Perception Reflected in Event-Related fMRI. *Cerebral Cortex* 17:2251–2257.

O’Connor K (2012) Auditory processing in autism spectrum disorder: A review. *Neuroscience & Biobehavioral Reviews* 36:836–854.

Ohala JJ (1975) The Temporal Regulation of Speech In Fant G, Tatham MAA, editors, *Auditory Analysis and Perception of Speech*, pp. 431–453. Academic Press.

Opitz A, Fox MD, Craddock RC, Colcombe S, Milham MP (2016) An integrated framework for targeting functional networks via transcranial magnetic stimulation. *NeuroImage* 127:86–96.

Ozimek E, Sek A (1988) AM Difference Limens for Noise Bands. *Acustica* 66:153–160.

Partridge A (2017) fMRI-Guided TMS to Auditory Cortex During Processing of Speech-Relevant Low-Level Stimuli. *Open Science Framework* .

Partridge AM (2016) fMRI-Guided State Dependent TMS of Auditory Cortex. *Open Science Framework* .

Pascual-Leone A, Walsh V, Rothwell J (2000) Transcranial magnetic stimulation in cognitive neuroscience – virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology* 10:232–237.

Patterson R, Nimmo-smith I, Holdsworth J, Rice P, Qoad M (1987) An Efficient Auditory Filterbank Based on the Gammatone Function In *Speech-Group Meeting of the Institute of Acoustcs on Auditory Modelling*, pp. 14–15.

Peelle JE (2014) Methodological challenges and solutions in auditory functional magnetic resonance imaging. *Frontiers in Neuroscience* 8.

Peelle JE, Davis MH (2012) Neural Oscillations Carry Speech Rhythm through to Comprehension. *Frontiers in Psychology* 3.

Peirce JW (2007) PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods* 162:8–13.

Pellicciari MC, Veniero D, Miniussi C (2017) Characterizing the Cortical Oscillatory Response to TMS Pulse. *Frontiers in Cellular Neuroscience* 11.

Pickering MJ, Garrod S (2013) An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences* 36:329–347.

Plewes DB, Kucharczyk W (2012) Physics of MRI: A primer. *Journal of Magnetic Resonance Imaging* 35:1038–1054.

Poeppel D (2003) The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication* 41:245–255.

Poeppel D (2014) The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology* 28:142–149.

Prendergast G, Green GGR (2012) Cross-channel amplitude sweeps are crucial to speech intelligibility. *Brain and Language* 120:406–411.

Prendergast G, Johnson SR, Green GGR (2010) Temporal dynamics of sinusoidal and non-sinusoidal amplitude modulation. *European Journal of Neuroscience* 32:1599–1607.

Prete G, Marzoli D, Brancucci A, Tommasi L (2016) Hearing it right: Evidence of hemispheric lateralization in auditory imagery. *Hearing Research* 332:80–86.

Price G, Mondinos N, Bose A (2015) Individualised Theta Burst stimulation using real-time feedback from ongoing EEG. *Brain Stimulation* 8:344–345.

Prins N (2012) The psychometric function: The lapse rate revisited. *Journal of Vision* 12:25–25.

Pulvermüller F, Huss M, Kherif F, Martin FMdP, Hauk O, Shtyrov Y (2006) Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences* 103:7865–7870.

R Core Team (2017) R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing.

Rammsayer T, Ulrich R (2012) The greater temporal acuity in the reminder task than in the 2AFC task is independent of standard duration and sensory modality. *Canadian Journal of Experimental Psychology* 66:26–31.

Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience* 2:79–87.

Ray P, Sinha VK, Tikka SK (2015) Adjuvant low-frequency rTMS in treating auditory hallucinations in recent-onset schizophrenia: A randomized controlled study investigating the effect of high-frequency priming stimulation. *Annals of General Psychiatry* 14:8.

Read JCA (2015) The place of human psychophysics in modern neuroscience. *Neuroscience* 296:116–129.

Rees A, Green GGR, Kay RH (1986) Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing Research* 23:123–133.

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

Riecke L, Formisano E, Herrmann CS, Sack AT (2015) 4-Hz Transcranial Alternating Current Stimulation Phase Modulates Hearing. *Brain Stimulation* 8:777–783.

Riecke L, Formisano E, Sorger B, Başkent D, Gaudrain E (2017) Neural Entrainment to Speech Modulates Speech Intelligibility. *Current Biology* 28:1–9.

Riesz RR (1928) Differential Intensity Sensitivity of the Ear for Pure Tones. *Physical Review* 31:867–875.

Rinderknecht MD, Ranzani R, Popp WL, Lambercy O, Gassert R (2018) Algorithm for improving psychophysical threshold estimates by detecting sustained inattention in experiments using PEST. *Attention, Perception, & Psychophysics* 80:1629–1645.

Rodenburg M (1977) Investigation of temporal effects with amplitude modulated signals. In Evans EF, Wilson JP, editors, *Psychophysics and Physiology of Hearing: An International Symposium.*, pp. 429–439. Academic Press, London.

Rohenkohl G, Cravo AM, Wyart V, Nobre AC (2012) Temporal Expectation Improves the Quality of Sensory Information. *Journal of Neuroscience* 32:8424–8428.

Romei V, Thut G, Silvanto J (2016) Information-Based Approaches of Noninvasive Transcranial Brain Stimulation. *Trends in Neurosciences* 39:782–795.

Ronconi L, Melcher D (2017) The Role of Oscillatory Phase in Determining the Temporal Organization of Perception: Evidence from Sensory Entrainment. *The Journal of Neuroscience* 37:10636–10644.

Ronconi L, Oosterhof NN, Bonmassar C, Melcher D (2017) Multiple oscillatory rhythms determine the temporal organization of perception. *Proceedings of the National Academy of Sciences* 114:13435–13440.

Rosanova M, Casali A, Bellina V, Resta F, Mariotti M, Massimini M (2009) Natural Frequencies of Human Corticothalamic Circuits. *Journal of Neuroscience* 29:7679–7685.

Rosen S (1992) Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 336:367–373.

Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America* 108:679–691.

Rossi S, Hallett M, Rossini PM, Pascual-Leone A (2009) Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology* 120:2008–2039.

Rounis E, Lee L, Siebner HR, Rowe JB, Friston KJ, Rothwell JC, Frackowiak RSJ (2005) Frequency specific changes in regional cerebral blood flow and motor system connectivity following rTMS to the primary motor cortex. *NeuroImage* 26:164–176.

Rufener KS, Oechslin MS, Zaehle T, Meyer M (2016) Transcranial Alternating Current Stimulation (tACS) differentially modulates speech perception in young and older adults. *Brain Stimulation* 9:560–565.

Rufener KS, Ruhnau P, Heinze HJ, Zaehle T (2017) Transcranial Random Noise Stimulation (tRNS) Shapes the Processing of Rapidly Changing Auditory Information. *Frontiers in Cellular Neuroscience* 11.

Rufener KS, Zaehle T, Oechslin MS, Meyer M (2016) 40Hz-Transcranial alternating current stimulation (tACS) selectively modulates speech perception. *International Journal of Psychophysiology* 101:18–24.

Ruff CC, Blankenburg F, Bjoertomt O, Bestmann S, Freeman E, Haynes JD, Rees G, Josephs O, Deichmann R, Driver J (2006) Concurrent TMS-fMRI and Psychophysics Reveal Frontal Influences on Human Retinotopic Visual Cortex. *Current Biology* 16:1479–1488.

Ruff CC, Driver J, Bestmann S (2009) Combining TMS and fMRI: From 'virtual lesions' to functional-network accounts of cognition. *Cortex* 45:1043–1049.

Saberi K, Perrott DR (1999) Cognitive restoration of reversed speech. *Nature* 398:760.

Sack AT, Cohen Kadosh R, Schuhmann T, Moerel M, Walsh V, Goebel R (2008) Optimizing Functional Accuracy of TMS in Cognitive Studies: A Comparison of Methods. *Journal of Cognitive Neuroscience* 21:207–221.

Saenz M, Langers DRM (2014) Tonotopic mapping of human auditory cortex. *Hearing Research* 307:42–52.

Santoro R, Moerel M, De Martino F, Goebel R, Ugurbil K, Yacoub E, Formisano E (2014) Encoding of Natural Sounds at Multiple Spectral and Temporal Resolutions in the Human Auditory Cortex. *PLoS Computational Biology* 10.

Saoud H, Josse G, Bertasi E, Truy E, Chait M, Giraud AL (2012) Brain–Speech Alignment Enhances Auditory Cortical Responses and Speech Perception. *Journal of Neuroscience* 32:275–281.

Sato M, Tremblay P, Gracco VL (2009) A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language* 111:1–7.

Schlittenlacher J, Moore BCJ (2016) Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing. *The Journal of the Acoustical Society of America* 140:3487–3495.

Schomers MR, Pulvermüller F (2016) Is the Sensorimotor Cortex Relevant for Speech Perception and Understanding? An Integrative Review. *Frontiers in Human Neuroscience* 10.

Schorer E (1986) Critical modulation frequency based on detection of AM versus FM tones. *The Journal of the Acoustical Society of America* 79:1054–1057.

Schremm A, Novén M, Horne M, Söderström P, van Westen D, Roll M (2018) Cortical thickness of planum temporale and pars opercularis in native language tone processing. *Brain and Language* 176:42–47.

Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences* 32:9–18.

Schütt HH, Harmeling S, Macke JH, Wichmann FA (2016) Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research* 122:105–123.

Schwarzbauer C, Davis MH, Rodd JM, Johnsrude I (2006) Interleaved silent steady state (ISSS) imaging: A new sparse imaging method applied to auditory fMRI. *NeuroImage* 29:774–782.

Scott SK, McGettigan C (2013) Do temporal processes underlie left hemisphere dominance in speech perception? *Brain and Language* 127:36–45.

Sek A, Moore BCJ (1994) The critical modulation frequency and its relationship to auditory filtering at low frequencies. *The Journal of the Acoustical Society of America* 95:2606–2615.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech Recognition with Primarily Temporal Cues. *Science* 270:303–304.

Sheft S, Yost WA (1990) Temporal integration in amplitude modulation detection. *The Journal of the Acoustical Society of America* 88:796–805.

Shen Y, Richards VM, Dai W (2015) A MATLAB toolbox for the efficient estimation of the psychometric function using the updated maximum-likelihood adaptive procedure. *Behavior research methods* 47:13–26.

Shen Y (2013) Comparing adaptive procedures for estimating the psychometric function for an auditory gap detection task. *Attention, perception & psychophysics* 75:771–780.

Shen Y (2014) Gap Detection and Temporal Modulation Transfer Function as Behavioral Estimates of Auditory Temporal Acuity Using Band-Limited Stimuli in Young and Older Adults. *Journal of speech, language, and hearing research* 57:2280–2292.

Shen Y, Richards VM (2012) A maximum-likelihood procedure for estimating psychometric functions: Thresholds, slopes, and lapses of attention. *The Journal of the Acoustical Society of America* 132:957–967.

Shen Y, Richards VM (2013) Temporal modulation transfer function for efficient assessment of auditory temporal resolution. *The Journal of the Acoustical Society of America* 133:1031–1042.

Silvanto J, Bona S, Marelli M, Cattaneo Z (2018) On the Mechanisms of Transcranial Magnetic Stimulation (TMS): How Brain State and Baseline Performance Level Determine Behavioral Effects of TMS. *Frontiers in Psychology* 9.

Silvanto J, Cattaneo Z (2017) Common framework for "virtual lesion" and state-dependent TMS: The facilitatory/suppressive range model of online TMS effects on behavior. *Brain and Cognition* 119:32–38.

Silvanto J, Muggleton NG (2008a) A novel approach for enhancing the functional specificity of TMS: Revealing the properties of distinct neural populations within the stimulated region. *Clinical Neurophysiology* 119:724–726.

Silvanto J, Muggleton NG (2008b) Testing the validity of the TMS state-dependency approach: Targeting functionally distinct motion-selective neural populations in visual areas V1/V2 and V5/MT+. *NeuroImage* 40:1841–1848.

Silvanto J, Muggleton NG, Cowey A, Walsh V (2007) Neural adaptation reveals state-dependent effects of transcranial magnetic stimulation. *European Journal of Neuroscience* 25:1874–1881.

Silvanto J, Pascual-Leone A (2008) State-Dependency of Transcranial Magnetic Stimulation. *Brain topography* 21:1–10.

Simpson AJR, Reiss JD, McAlpine D (2013) Tuning of Human Modulation Filters Is Carrier-Frequency Dependent. *PLOS ONE* 8:e73590.

Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America* 114:3394–3411.

Smith NA, Trainor LJ, Gray K, Plantinga JA, Shore DI (2008) Stimulus, Task, and Learning Effects on Measures of Temporal Resolution: Implications for Predictors of Language Outcome. *Journal of Speech, Language, and Hearing Research* 51:1630–1642.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23:S208–S219.

Sowman PF, Flavel SC, McShane CL, Sakuma S, Miles TS, Nordstrom MA (2009) Asymmetric Activation of Motor Cortex Controlling Human Anterior Digastric Muscles During Speech and Target-Directed Jaw Movements. *Journal of Neurophysiology* 102:159–166.

Spratling MW (2017) A review of predictive coding algorithms. *Brain and Cognition* 112:92–97.

Stokes MG, Chambers CD, Gould IC, English T, McNaught E, McDonald O, Mattingley JB (2007) Distance-adjusted motor threshold for transcranial magnetic stimulation. *Clinical Neurophysiology* 118:1617–1625.

Stone MA, Moore BCJ (2014) Amplitude-modulation detection by recreational-noise-exposed humans with near-normal hearing thresholds and its medium-term progression. *Hearing Research* 317:50–62.

Stone MA, Moore BCJ, Greenish H (2008) Discrimination of envelope statistics reveals evidence of sub-clinical hearing damage in a noise-exposed population with 'normal' hearing thresholds. *International Journal of Audiology* 47:737–750.

Suppa A, Huang YZ, Funke K, Ridding MC, Cheeran B, Di Lazzaro V, Ziemann U, Rothwell JC (2016) Ten Years of Theta Burst Stimulation in Humans: Established Knowledge, Unknowns and Prospects. *Brain Stimulation* 9:323–335.

Talavage TM, Hall DA (2012) How challenges in auditory fMRI led to general advancements for the field. *NeuroImage* 62:641–647.

Tang H, Crain S, Johnson BW (2016) Dual temporal encoding mechanisms in human auditory cortex: Evidence from MEG and EEG. *NeuroImage* 128:32–43.

Tansley BW, Suffield JB (1983) Time course of adaptation and recovery of channels selectively sensitive to frequency and amplitude modulation. *The Journal of the Acoustical Society of America* 74:765–775.

Teng X, Tian X, Rowland J, Poeppel D (2017) Concurrent temporal channels for auditory processing: Oscillatory neural entrainment reveals segregation of function at different scales. *PLOS Biology* 15:e2000812.

Terney D, Chaieb L, Moliadze V, Antal A, Paulus W (2008) Increasing Human Brain Excitability by Transcranial High-Frequency Random Noise Stimulation. *Journal of Neuroscience* 28:14147–14155.

Thielscher A (2013) IS 3. TMS field modelling-status and next steps. *Clinical Neurophysiology* 124:e40.

Thielscher A, Antunes A, Saturnino GB (2015) Field modeling for transcranial magnetic stimulation: A useful tool to understand the physiological effects of TMS? In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 222–225.

Thut G, Bergmann TO, Fröhlich F, Soekadar SR, Brittain JS, Valero-Cabré A, Sack AT, Miniussi C, Antal A, Siebner HR, Ziemann U, Herrmann CS (2017) Guiding transcranial brain stimulation by EEG/MEG to interact with ongoing brain activity and associated functions: A position paper. *Clinical Neurophysiology* 128:843–857.

Thut G, Schyns PG, Gross J (2011) Entrainment of Perceptually Relevant Brain Oscillations by Non-Invasive Rhythmic Stimulation of the Human Brain. *Frontiers in Psychology* 2.

Thut G, Veniero D, Romei V, Miniussi C, Schyns P, Gross J (2011) Rhythmic TMS Causes Local Entrainment of Natural Oscillatory Signatures. *Current Biology* 21:1176–1185.

Treutwein B (1995) Adaptive psychophysical procedures. *Vision Research* 35:2503–2522.

Ueda K, Nakajima Y, Ellermeier W, Kattner F (2017) Intelligibility of locally time-reversed speech: A multilingual comparison. *Scientific Reports* 7:1782.

Ulrich R, Vorberg D (2009) Estimating the difference limen in 2AFC tasks: Pitfalls and improved estimators. *Attention, Perception, & Psychophysics* 71:1219–1227.

Van Der Werf YD, Paus T (2006) The neural response to transcranial magnetic stimulation of the human motor cortex. I. Intracortical and cortico-cortical contributions. *Experimental Brain Research* 175:231–245.

van Ede F, Quinn AJ, Woolrich MW, Nobre AC (2018) Neural Oscillations: Sustained Rhythms or Transient Burst-Events? *Trends in Neurosciences* 41:415–417.

Verhulst S, Altoè A, Vasilkov V (2018) Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss. *Hearing Research* 360:55–75.

Vernet M, Bashir S, Yoo WK, Oberman L, Mizrahi I, Ifert-Miller F, Beck CJ, Pascual-Leone A (2014) Reproducibility of the effects of theta burst stimulation on motor cortical plasticity in healthy participants. *Clinical Neurophysiology* 125:320–326.

Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. *The Journal of the Acoustical Society of America* 66:1364–1380.

Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S (2011) Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences of the United States of America* 108:20754–20759.

Wagner T, Rushmore J, Eden U, Valero-Cabre A (2009) Biophysical foundations underlying TMS: Setting the stage for an effective use of neurostimulation in the cognitive neurosciences. *Cortex* 45:1025–1034.

Wakefield GH, Viemeister NF (1990) Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise. *The Journal of the Acoustical Society of America* 88:1367–1373.

Wallis SA, Baker DH, Meese TS, Georgeson MA (2013) The slope of the psychometric function and non-stationarity of thresholds in spatiotemporal contrast vision. *Vision Research* 76:1–10.

Wang X (2018) Cortical Coding of Auditory Features. *Annual Review of Neuroscience* 41:527–552.

Wang X, Lu T, Liang L (2003) Cortical processing of temporal modulations. *Speech Communication* 41:107–121.

Watkins KE, Strafella AP, Paus T (2003) Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–994.

Watson AB (2017) QUEST+: A general multidimensional Bayesian adaptive psychometric method. *Journal of Vision* 17:10–10.

Watson AB, Pelli DG (1983) Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics* 33:113–120.

Weisz N, Lithari C (2017) Amplitude modulation rate dependent topographic organization of the auditory steady-state response in human auditory cortex. *Hearing Research* 354:102–108.

Weisz N, Steidle L, Lorenz I (2012) Formerly known as inhibitory: Effects of 1-Hz rTMS on auditory cortex are state-dependent. *European Journal of Neuroscience* 36:2077–2087.

Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics* 63:1293–1313.

Wilsch A, Neuling T, Obleser J, Herrmann CS (2018) Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. *NeuroImage* 172:766–774.

Wilson M, Knoblich G (2005) The case for motor involvement in perceiving conspecifics. *Psychological Bulletin* 131:460–473.

Witton C, Talcott JB, Henning GB (2017) Psychophysical measurements in children: Challenges, pitfalls, and considerations. *PeerJ* 5:e3231.

Wojtczak M, Viemeister NF (2003) Suprathreshold effects of adaptation produced by amplitude modulation. *The Journal of the Acoustical Society of America* 114:991–997.

Wojtczak M, Viemeister NF (2005) Forward masking of amplitude modulation: Basic characteristics. *The Journal of the Acoustical Society of America* 118:3198–3210.

Worsley KJ (2001) Statistical Analysis of Activation Images In *Functional MRI: An Introduction to Methods*. Oxford University Press, Oxford, UK.

Xiang J, Poeppel D, Simon JZ (2013) Physiological evidence for auditory modulation filterbanks: Cortical responses to concurrent modulations. *The Journal of the Acoustical Society of America* 133:EL7–EL12.

Young ED (2008) Neural representation of spectral and temporal information in speech. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 363:923–945.

Zaehle T, Wüstenberg T, Meyer M, Jäncke L (2004) Evidence for rapid auditory perception as the foundation of speech processing: A sparse temporal sampling fMRI study. *European Journal of Neuroscience* 20:2447–2456.

Zatorre RJ, Belin P (2001) Spectral and Temporal Processing in Human Auditory Cortex. *Cerebral Cortex* 11:946–953.

Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences* 6:37–46.

Zchaluk K, Foster DH (2009) Model-free estimation of the psychometric function. *Attention, Perception, & Psychophysics* 71:1414–1425.

Zoefel B, Archer-Boyd A, Davis MH (2018) Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Current Biology* 28:401–408.

Zoefel B, Davis MH (2017) Transcranial electric stimulation for the investigation of speech perception and comprehension. *Language, Cognition and Neuroscience* 32:910–923.

Zoefel B, ten Oever S, Sack AT (2018) The Involvement of Endogenous Neural Oscillations in the Processing of Rhythmic Input: More Than a Regular Repetition of Evoked Neural Responses. *Frontiers in Neuroscience* 12.

Zoefel B, VanRullen R (2016) EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage* 124:16–23.

Zwicker E (1952) Audibility limits for amplitude and frequency modulation of a tone. *Acustica* 2:AB125 – AB133.