Preparatory and selective attention during multi-talker listening in normal and impaired hearing

Emma Holmes

PhD

University of York

Psychology

October 2014

Abstract

One of the great challenges of hearing research is to work out how listeners can perceive what one talker is saying when other talkers are speaking at the same time. Faced with this requirement for 'multi-talker listening', normally-hearing listeners achieve improved speech intelligibility when they know characteristics of an upcoming talker before he or she begins to speak. One aim was to investigate the time course of this improvement in intelligibility and the brain activity that accompanies it. A task was devised in which participants had to report key words spoken by a 'target' talker when one or two other talkers spoke simultaneously. Before the talkers began to speak, a visual cue indicated the location (left/right) or gender (male/female) of the target talker. The accuracy and latency of reporting key words progressively improved when participants had longer to prepare for the location or gender of the target talker. Preparatory brain activity, measured with electro-encephalography, began with a short latency (< 100 ms) after the reveal of the visual cue and was sustained until the talkers began to speak.

Hearing-impaired listeners, both children and adults, typically show poorer speech intelligibility during multi-talker listening than normally-hearing listeners. One advantage of the experimental design was that brain activity during preparatory attention (before the onset of acoustical stimuli) could be compared between normally-hearing and hearing-impaired listeners and atypical attention identified, without confounding differences in transduction at the auditory periphery. This thesis demonstrates atypical preparatory EEG activity in children, aged 7-16 years, with bilateral moderate cochlear hearing loss, which provides evidence for atypical preparatory attention. Therefore, atypical preparatory attention might be one factor that contributes to poorer speech intelligibility in noisy environments. An implication is that acoustic hearing aids may not have the potential alone to restore normal processing of acoustical stimuli in hearing-impaired listeners.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my supervisor, Quentin Summerfield, for the opportunity to carry out this project. I greatly appreciate the mentorship, support, and guidance that he has provided throughout the project, including comments on earlier drafts of the chapters that constitute this thesis. I would also like to thank Pádraig Kitterick, who contributed to the design of these experiments in the initial stages of the project and helped it to get off the ground. Since then, we have had many interesting conversations about the methods and results that are reported in this thesis.

The members of my Thesis Advisory Panel—Gary Green and Emma Hayiou-Thomas—provided useful feedback at regular intervals throughout the project. I appreciate their ideas and constructive feedback on the work that I presented.

For the experiments that recruited hearing-impaired children, I greatly appreciate the time and effort of a number of clinicians who sent letters to potential participants or who helped with this process. They include Gerard Reilly and Kate Iley of the York Teaching Hospital NHS Foundation Trust, Aung Nyunt of Hull and East Yorkshire Hospitals NHS Trust, Sanjay Verma and Mirriam Iqbal of The Leeds Teaching Hospitals NHS Trust, and Christopher Raine, Rob Gardner, and Sara Morgan of Bradford Teaching Hospitals NHS Foundation Trust. I would also like to thank Andrew Coatesworth for helping to ignite these collaborations.

Importantly, I could not have completed this thesis without the support of my friends and family. They have kept my spirits high and have encouraged a positive work-life balance.

Financial Support

The first year of this project was funded by the Department of Psychology of the University of York. The second and third years were funded by the Goodricke Appeal Fund—a charity for hearing-impaired children. I would like to thank both of these sources for providing generous PhD studentships, both of which have allowed me to carry out this project.

I am also grateful for a number of travel awards, which have allowed me to attend both UK and international conferences. Partial funding was provided by the Department of Psychology of the University of York and the Goodricke Appeal Fund. Additional funding was provided by Deafness Research UK (February 2013), Guarantors of Brain (February 2013), Association for Research in Otolaryngology (February 2013), Experimental Psychology Society (January 2014 and February 2014), and Action on Hearing Loss (February 2014). The opportunity to attend these conferences has been extremely valuable and I am very grateful for the support provided by these awards.

Declaration

The candidate is the sole author of this thesis. This thesis comprises original work that has not previously been presented for an award at this, or any other, University. All experiments were designed by the candidate with assistance from the supervisor, Professor Quentin Summerfield, and Dr. Padraig Kitterick. All testing and analyses were conducted by the candidate. All sources are acknowledged as references.

Conference presentations and proceedings

Holmes, E., Kitterick, P. T. & Summerfield, A. Q. (2014). Do children with hearing loss show atypical attention during 'cocktail party' listening? Paper presented at British Association for Cognitive Neuroscience Annual Scientific Meeting, York, UK.

Holmes, E., Kitterick, P. T. & Summerfield, A. Q. (2014). Do children with hearing loss show atypical attention during 'cocktail party' listening? Paper presented at British Society for Audiology 5th Annual Conference and Experimental and Clinical Papers Short Meeting, Keele, UK.

Holmes, E., Kitterick, P. T. & Summerfield, A. Q. (2014). Timing and Specificity of Preparatory Attention in Cocktail-Party Listening. Paper presented at the Association for Research in Otolaryngology 37th Mid-Winter Meeting, San Diego, CA, USA.

Holmes, E., Kitterick, P. T. & Summerfield, A. Q. (2014). Exploring the benefit of preparatory auditory attention in multi-talker listening: Evidence from EEG. Paper presented at Experimental Psychology Society Meeting, University College London, UK.

Holmes, E., Kitterick, P. T. & Summerfield, A. Q. (2013). Preparatory attention in cocktail party listening. Paper presented at the British Society for Audiology 4th Annual Conference and Experimental and Clinical Papers Short Meeting, Keele, UK.

Holmes, E., Kitterick, P. T. & Summerfield, A. Q. (2013). Preparatory and selective attention in children during multi-talker listening. Paper presented at the Association for Research in Otolaryngology 36th Mid-Winter Meeting, Baltimore, MD, USA.

Holmes, E., Kitterick, P. T. & Summerfield, A. Q. (2012). Distinguishing preparatory from selective attention during multi-talker listening. Paper presented at the British

Society for Audiology 3rd Annual Conference and Experimental and Clinical Papers Short Meeting, Nottingham, UK.

Holmes, E., & Summerfield, A. Q. (2012). Auditory attention during multi-talker listening: Evidence from event-related potentials. Paper presented at the White Rose DTC Annual Spring Conference, University of York, UK.

Overview of Thesis

Speech is typically encountered in the presence of other sounds, including the voices of other talkers. The ability to identify words spoken by one talker when other talkers are speaking is sometimes referred to as 'multi-talker listening' or 'the cocktail party problem'. In this situation, the control of auditory attention is critical for successful communication. One key finding is that normally-hearing listeners show improved speech intelligibility during multi-talker listening when they know attributes of a talker before he or she begins to speak (e.g. Best, Marrone, Mason, Kidd, & Shinn-Cunningham, 2009; Kitterick, Bailey, & Summerfield, 2010). However, the mechanisms that underlie this improvement are not fully understood. A better understanding of the mechanisms that improve speech intelligibility for normally-hearing listeners has the potential to improve understanding of the processes that contribute to poorer speech intelligibility in hearing-impaired listeners.

Overall, this thesis aimed to investigate the mechanisms by which participants prepare their attention when they know attributes of a talker before he or she begins to speak and the mechanisms by which participants attend selectively to a talker while multiple talkers speak simultaneously. This thesis examined these two processes—hereafter referred to as 'preparatory attention' and 'selective attention'—in normally-hearing and hearing-impaired listeners.

Multi-talker listening is particularly challenging for listeners with impaired hearing (e.g. Dubno, Dirks, & Morgan, 1984; Helfer & Freyman, 2008). To investigate preparatory and selective attention in hearing-impaired listeners, the experiments in this thesis focused on children, aged between 7 and 16 years, who had moderate cochlear hearing loss, despite the fact that most previous experiments have typically studied older adults with hearing loss. The rationale was that, in older adults, the central consequences of hearing loss are difficult to separate from general cognitive decline with older age that is independent from (although perhaps correlated with) peripheral hearing loss. A previous experiment found that normally-hearing children aged 10–15 years, like adults, can benefit from advance cueing in noisy environments (Dhamani, Leung, Carlile, & Sharma, 2013). Therefore, atypical attention during multi-talker listening in hearing-impaired children, compared to normally-hearing children of the same age, can be attributed to peripheral hearing loss.

Chapter 1 summarises previous research that has improved understanding of attention, factors that influence the accuracy of speech intelligibility during multi-talker listening, and possible processes that underlie poorer performance in hearing-impaired listeners. Two key experiments previously investigated the brain regions that were active in normally-hearing listeners during multi-talker listening using functional magnetic resonance imaging (fMRI[1]; Hill & Miller, 2010) and magnetoencephalography (MEG; Lee et al., 2013). In the two previous experiments, participants were cued to attend to the location or fundamental frequency of an upcoming talker and their brain activity was measured during preparatory and selective attention. These experiments identified *regions of the brain* that showed significant activity during multi-talker listening. However, the experiments were not designed to illuminate the *timing* of brain activity during preparatory and selective attention.

To investigate the timing of brain activity, the experiments in this thesis used electroencephalography (EEG)—a technique that measures electrical neural activity directly and non-invasively from the scalp with high temporal resolution. **Chapter 2** discusses the neural basis of EEG activity and introduces the analyses that are employed in the EEG experiments reported in this thesis.

The experiments reported in Chapters 3–5 had two main aims: (1) to devise a technique for measuring preparatory and selective attention during multi-talker listening in normally-hearing listeners, which would also be suitable for normally-hearing and hearing-impaired children, and (2) to examine the time-course of preparatory and selective attention in normally-hearing listeners. **Chapter 3** reports three experiments that examined the time-course of EEG activity during a two-talker listening task. Two experiments involved normally-hearing adults and one experiment involved normally-hearing children aged 7–13 years. The experiments reported in **Chapter 4** investigated how the duration of preparation time affects speech intelligibility. Specifically, the experiments aimed to distinguish between the hypotheses that there is a 'threshold' of time necessary for successful preparation, or that increasing durations of preparation time produce progressive improvements in speech intelligibility. The experiments employed a task in which participants were cued to the location or gender of a target talker who spoke in a mixture of two or three talkers. **Chapter 5** reports an experiment that measured EEG activity in normally-hearing adults during a three-talker listening task. Overall, Chapters 3–5 provide detailed information about the time course of preparatory attention in normally-hearing listeners. Based on these results, the three-talker task was selected to measure preparatory and selective attention in normally-hearing and hearing-impaired children, the results of which are reported in Chapter 6.

---

[1] The chapters in this thesis were written with the intention that they could stand alone. Therefore, abbreviations are redefined at their first instance in each chapter.

One previous experiment showed less improvement in the accuracy of speech intelligibility in hearing-impaired than normally-hearing listeners when they received information about the spatial location of an upcoming talker (Best et al., 2009). This result is consistent with the idea that hearing-impaired listeners show atypical preparatory attention for multi-talker listening. One possibility is that hearing-impaired listeners do not deploy preparatory attention to the same extent as normally-hearing listeners. An alternative is that hearing-impaired listeners need more time to prepare effectively for an upcoming talker than do normally-hearing listeners. The experiments reported in **Chapter 6** investigated preparatory attention in children with moderate hearing loss of cochlear origin. Hearing-impaired children and a comparison group of normally-hearing children participated. First, this chapter reports experiments that investigated whether hearing-impaired children achieved better speech intelligibility when they were cued to the location or gender of an upcoming talker than when they did not know this information in advance. The final two experiments investigated whether hearing-impaired children show atypical EEG activity during multi-talker listening compared to normally-hearing children.

Chapter 7 provides a summary of the results from Chapters 3–6. This chapter also discusses implications of the results and key issues raised in this thesis. It concludes by highlighting directions for future research.

Chapter 1

Auditory Attention and Hearing Loss

Auditory attention is critical for successful communication because speech is typically encountered in the presence of other sounds, including the voices of other talkers. Listeners must flexibly allocate attentional resources to a talker of interest and ignore distracting sounds. The ability to identify words spoken by one talker when other talkers are speaking is sometimes referred to as 'multi-talker listening' or 'the cocktail party problem'.

Multi-talker listening occurs so frequently in everyday life that normally-hearing listeners may take it for granted (Münte, Spring, Szycik, & Noesselt, 2010). At positive signal-to-noise ratios (SNRs; i.e. where the speech of interest occurs with greater energy than the speech of competing talkers), normally-hearing listeners do not have difficulty identifying words spoken by one talker. However, even normally-hearing listeners experience difficulty when the competing speech occurs with greater energy than the target speech. The accuracy of reporting words spoken by a target talker decreases as the SNR decreases from 0 to -12 decibels (dB; e.g. Freyman, Balakrishnan, & Helfer, 2001; Rosen, Souza, Ekelund, & Majeed, 2013). Understanding the factors that improve speech intelligibility is becoming increasingly important in a technologically advancing age, where high demands are placed upon listeners' perceptual systems due to the growing popularity of mobile phones, satellite navigation systems, and portable music devices.

For listeners with hearing loss, identifying speech can be a struggle even at positive SNRs. Duquesnoy (1983) found that speech reception thresholds (SRTs; i.e. the lowest SNRs required to correctly identify 50% of target sentences) were more than 10 dB SNR higher in hearing-impaired listeners than in normally-hearing listeners. Although speech recognition in quiet can be improved by an acoustic hearing aid, listening in background noise often remains difficult (Marrone, Mason, & Kidd, 2008a). The central consequences of hearing loss, including how hearing-impaired listeners direct attention to speech in noisy environments, are not fully understood. A greater understanding of the factors that benefit the accuracy of speech intelligibility for normally-hearing listeners might help to better understand the difficulties faced by hearing-impaired listeners.

This chapter reviews previous research on multi-talker listening. The focus of the first half is normally-hearing listeners and the second half is hearing-impaired listeners. I begin by discussing attention to visual stimuli and to auditory non-speech stimuli, since this research has led to greater understanding of the brain regions underlying attention—of which many parallels can be drawn between visual and auditory modalities. Next, I discuss the factors that affect speech intelligibility in noise. Towards the end of the chapter, I discuss the rationale for expecting atypical attentional processing of speech in hearing-impaired listeners. The chapter will conclude by highlighting the key questions that this thesis addresses.

1.1. Attention within and between modalities

1.1.1. Advantages of attention

1.1.1.1. Perceptual benefits

Many researchers view attention as a result of a limited capacity for processing information, acting as a 'filter' on the perceptual system (Kahneman, 1973; Norman & Bobrow, 1975). However, directing attention to a stimulus of interest has advantages beyond simply 'filtering' the incoming information. For example, directing attention to a particular location in space has consistently been found to improve the accuracy and latency with which a stimulus is detected at that location, across visual (Coull & Nobre, 1998; Posner, Petersen, Fox, & Raichle, 1988; Posner, Snyder, & Davidson, 1980) and auditory (Arbogast & Kidd, 2000; Hink, Van Voorhis, Hillyard, & Smith, 1977; Woods, Hillyard, & Hansen, 1984) modalities, as well as for multisensory stimuli (Spence & Parise, 2010; Zampini et al., 2005).

1.1.1.2. Enhanced neural responses

The perceptual benefits of attended stimuli are associated with enhanced neural responses in areas representing attended stimulus dimensions (Shamma & Micheyl, 2010; Wild et al., 2012). Several experiments have measured neural responses using electroencephalography (EEG) when participants attend to tone sequences presented to one ear and ignore tones presented to the opposite ear. Attended tones evoke significantly larger potentials than unattended tones. This difference has been observed over central and frontal electrodes as early as 20 ms after the onset of the tone and also at longer latencies (60–150 ms) over central electrodes (Woldorff, Hansen, & Hillyard, 1987; Woldorff & Hillyard, 1991). The short-latency differences observed in these experiments are thought to originate from primary auditory cortex (A1; Woldorff et al., 1993). Greater responses for attended over unattended stimuli have also been observed using magnetoencephalography (MEG; Woldorff et al., 1993) and functional magnetic resonance imaging (fMRI; Woodruff et al., 1996). These modulations of neural responses are thought to underlie the perceptual benefit of attending to a particular spatial location.

1.1.2. Types of attention

Allport (1993) argued that attention does not refer to a single process, but rather to a set of phenomena. In everyday situations, different aspects of attention typically operate together. Duncan (2006) noted that observers often attend to aspects of a scene as a coherent whole, including information from the senses, working memory, and their goals and actions. Nevertheless, several distinctions have been made between different types of attentional process. This section will discuss three distinctions: exogenous and endogenous attention, object- and feature-based attention, and selective and divided attention.

1.1.2.1. Exogenous and endogenous attention

Top-down attention can be flexibly directed to a stimulus of interest, which enables an observer to select a stimulus of interest in accordance with his or her goals. However, some stimuli are particularly salient among others and can grab attention relatively automatically (Desimone & Duncan, 1995; Yantis, 2005). Stimulus-driven attention is typically referred to as 'exogenous' attention, while strategic, goal-driven attention is referred to as 'endogenous' attention.

There are several functional differences between exogenous and endogenous attention, suggesting that they are different processes. For example, Jonides (1981) demonstrated that visual endogenous attention is disrupted by a secondary verbal-memory task, while exogenous attention is not. Arrow stimuli were used as visual cues in both conditions, but endogenous cues were presented centrally, whereas exogenous cues were presented at the position of the upcoming target stimulus. Participants were asked to recall three, five, or seven digits at the end of each trial. Memory load affected reaction times (RTs) in the endogenous condition, but not the exogenous condition. This result suggests that endogenous attention relies on verbal working memory, while exogenous attention does not.

In addition to functional differences between exogenous and endogenous attention, neuroimaging studies suggest distinct neural substrates. Endogenous attention relies on frontal and parietal areas, while exogenous attention involves a right-hemisphere ventral network (Corbetta, Patel, & Shulman, 2008).

The remainder of this thesis will primarily focus on endogenous attention, for which the underlying mechanisms will be described later in this chapter.

1.1.2.2. Object-, space-, and feature-based attention

Over the past decade, researchers have debated whether attention is object-based, space-based, or feature-based (e.g. Alain & Arnott, 2000). These hypotheses differ in the stimulus attributes they propose that attention operates upon. Object-based accounts assume that attention operates after location and feature information are processed and after the presented stimuli are analysed semantically. Whereas, space- and feature-based accounts assume that attention operates upon either spatial locations or features and that unattended stimuli do not undergo further processing (e.g. semantic processing). Driver (2001) suggested that this debate is a recent reinterpretation of the long-lasting debate between early and late selection.

Early selection refers to a perceptual filter of sensory information. This account assumes that attention operates at an early stage of processing and only attended stimuli are processed (Broadbent, 1958). Whereas, late selection assumes that the filter operates later in processing, at the response selection stage, so that even unattended stimuli are processed (Deutsch & Deutsch, 1963). Lavie and Tsal (1994) reconciled much of the previous evidence by proposing that perceptual load determines the stage at which selection takes place. Processing of unattended stimuli is only prevented when the load of relevant information is sufficiently high to demand all of the available resources. Under low loads, Lavie (2001, 2005) argued that distracters are automatically processed, but not acted upon. Lavie and Tsal noted that results apparently favouring late selection had typically been obtained when the load was low, whereas results favouring early selection were obtained under high load. These observations are consistent with the perceptual load hypothesis, and many further experiments aimed at testing perceptual load theory are also consistent with this hypothesis (see Lavie, 2005 for a review).

Taking perceptual load theory into account, it is possible that load determines whether object-, space-, or feature-based attention occurs. One hypothesis is that attention operates on objects under low load, but on space or features under high load. The rationale is that low perceptual loads enable observers to extract location and feature information from the presented stimuli, whereas high perceptual loads are demanding, meaning that observers may extract only feature *or* location information. This hypothesis has not been tested and many researchers have assumed that attention is either object-based or feature-based and does not change under different levels of perceptual load (e.g. Scholl, 2001; Shinn-Cunningham, 2008).

1.1.2.3. Selective and divided attention

The distinction between selective and divided attention may also be linked to early and late selection, respectively. Selective attention is characterised by monitoring one stimulus at the expense of others, whereas divided attention is characterised by monitoring multiple inputs (Shafiro & Gygi, 2007). Lavie and Tsal (1994) proposed that irrelevant processing is only prevented when the perceptual load is high. Therefore, selective attention is more likely to operate under high load (since the attended stimulus demands all of the available resources), while divided attention is more likely to operate under low load (since more than one stimulus may undergo further processing). Naatanen (1990) proposed that low stimulus presentation rates allow participants to monitor multiple stimuli should they wish to do so, even if monitoring multiple stimuli does not improve their performance on the current task. Importantly, even if an experiment aims to measure selective or divided attention, load conditions should be taken into account because they may have unintended consequences for the number of stimuli that are monitored.

Rather than being discrete aspects of attention, as is traditionally assumed, attention may vary on a continuum between selective and divided attention and can even vary within a trial. For example, in certain tasks, participants must first monitor all stimuli to determine which is the target, then selectively attend to the target to determine the correct response (e.g. Kitterick et al., 2010; Shafiro & Gygi, 2007). When reviewing studies of selective and divided attention, it is important to consider this distinction.

1.1.3. Neural basis of attention

There are thought to be many similarities between visual and auditory endogenous attention (Larson & Lee, 2013b; Shinn-Cunningham, 2008; Shomstein & Yantis, 2006). Therefore, knowledge of the neural bases of visual attention can inform hypotheses about auditory attention.

1.1.3.1. Visual attention

It is widely accepted that endogenous visual attention involves a fronto-parietal network of brain activity. Lesion studies show that the parietal cortex is critical for spatial attention (Posner, Inhoff, Friedrich, & Cohen, 1987). Converging evidence from neuroimaging shows activation of frontal and parietal regions during both endogenous spatial (Corbetta et al., 2005; Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Nobre et al., 1997; Yantis et al., 2002) and endogenous non-spatial (Giesbrecht, Woldorff, Song, & Mangun, 2003; Shulman, d'Avossa, Tansy, & Corbetta, 2002) attention. Activity that occurs in similar brain regions when observers attend to spatial and non-spatial attributes of a stimulus is often referred to as 'domain-general' attentional processing. Areas that contribute to domain-general activity include the intraparietal sulcus and posterior and medial frontal cortex (Slagter, Giesbrecht, & Kok, 2007).

In addition to fronto-parietal regions, activity is observed in areas of sensory cortex that are specialised for processing the attended dimension. For example, Chawla, Rees, and Friston (1999) asked participants to attend to the colour or motion of a random dot motion stimulus in a delayed match-to-sample task. In the motion condition, participants had to discriminate between fast- and slow-moving dots. In the colour condition, participants had to detect dots of a different (slightly pinker) colour. Attending to motion produced greater hemodynamic responses in area V5 than attending to colour. Conversely, attending to colour preferentially activated area V4. Since these areas are specialised for processing motion and colour respectively (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Conway, 2009), the results provide evidence that attention increases activity in regions specialised for processing the attended dimension.

The ‘biased competition’ model of attention attempts to explain the enhanced neural response for attended stimuli. This account proposes that stimuli compete for representation in sensory cortices. Furthermore, biased competition assumes that attending to a spatial location leads to greater neural responses for stimuli presented at attended, than unattended, locations (Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999). Single-unit recordings provide evidence that is consistent with biased competition in visual cortices. For example, Luck, Chelazzi, Hillyard, and Desimone (1997) recorded from cells in V1, V2, and V4 of macaque monkeys. The monkeys were shown sequences of squares and rectangles and had to detect the presence of squares at an attended location. The researchers compared responses to attended and unattended stimuli. Stimuli that were subsequently presented as attended and unattended stimuli both evoked rapid firing rates when they were presented in the cell’s receptive field (RF) alone. However, when the stimuli were presented simultaneously at different locations within the cell’s RF, the response to the unattended stimulus was reduced. This result demonstrates that: (1) stimuli compete for representation in visual cortices (if stimuli did not compete, then both stimuli should produce similar responses when presented simultaneously); and (2) top-down attention influences which stimulus receives preferential representation. Together, the results of the aforementioned studies are consistent with the hypothesis that attention activates a domain-general fronto-parietal network and influences the magnitude of neural responses in areas that are specialised for processing the attended dimension.

In addition to modulating response magnitude, selective visual attention affects neural synchrony. Fries (2001) reported that cortical regions specialised for processing the attended dimension displayed increased gamma-band (35–90 Hz) and decreased alpha-band (8–12 Hz) synchrony. Synchronous activity is a potential mechanism by which brain regions communicate. This idea results from the logic that neurons have limited temporal integration windows and, thus, synchronous firing is likely to have a larger impact on downstream neurons than asynchronous firing. It has been suggested that neural synchrony in the gamma band during selective attention tasks might indicate a flow of activity from prefrontal to sensory regions. For example, Gregoriou, Gotts, Zhou, and Desimone (2009) recorded multi-unit responses and local field potentials from the frontal eye fields (FEF) and area V4. Monkeys were trained to attend to one of three visual stimuli and release a bar when the attended stimulus changed colour. Firing rates in FEF and V4 both increased when attention was directed inside the area's RF than when attention was directed outside the RF. However, the latency of increased firing occurred earlier in FEF than in V4. Furthermore, attention increased gamma-band coherence (a measure of phase locking) between FEF and V4. Granger causality analysis (a method based on linear regression) indicated that early gamma-band local field potentials in FEF (110 ms) were predictive of local field potentials in V4. This result led the researchers to conclude that attentional gamma-band activity in FEF affects neural synchrony in area V4. However, the precise mechanism by which attention affects synchronous activity is not fully understood.

1.1.3.2. Auditory attention

Evidence suggests that a fronto-parietal network also controls auditory attention. Patients with parietal lobe damage often experience neglect in audition as well as in vision (Bellmann, Meuli, & Clarke, 2001; Pavani, Làdavas, & Driver, 2003; Spierer, Meuli, & Clarke, 2007), which is consistent with the idea that the networks for auditory and visual attention might at least partially overlap. Auditory neuroimaging also supports a fronto-parietal network for attention. For example, Salmi, Rinne, Koistinen, Salonen, and Alho (2009) presented participants with two streams of band-limited noise bursts through headphones—one to the left and one to the right ear. They cued the target stream with a central visual cue comprised of a red and green arrow. For half of participants, the *red* arrows denoted the target location (left/right), while for the other half, the *green* arrows denoted the target location. Participants' task was to detect shorter-duration tones within the target stream. The results revealed activity in regions of the fronto-parietal network, including superior parietal lobule, intraparietal sulcus, FEF, and inferior and medial frontal gyri.

Attending to acoustical stimuli also activates regions of auditory cortex specialised for processing the attended dimension. Several studies have contrasted blood-oxygen level-dependent (BOLD) activity when participants attend to auditory or visual stimuli.

Importantly, both visual and acoustical stimuli are presented in all conditions. These studies observe a similar fronto-parietal network during attend-visual and attend-auditory conditions, but a double dissociation in regions specialised for visual and auditory processing: greater activity in auditory areas when acoustical stimuli are attended and greater activity in visual areas when visual stimuli are attended (Kawashima et al., 1999; Sabri et al., 2008; Salmi, Rinne, Degerman, & Alho, 2007; Wild et al., 2012). Kawashima *et al.* (1999) propose that these results support a modality-dependent selective attention mechanism. However, although the specific regions differ when participants are attending to vision or audition, the pattern of activity appears to be the same in both cases: activity in a domain-general fronto-parietal network and in regions specialised for processing the attended dimension. When attending to visual stimuli, the specialised regions lie in visual cortices, but when attending to acoustical stimuli, the specialised regions lie in auditory cortices.

Taking the argument one step further, when attending to different attributes of acoustical stimuli within the auditory modality, differences in brain activity are observed in areas specialised for processing these attributes. For example, Ahveninen *et al.* (2006) presented Finnish vowel sounds from two possible locations: 0 degrees azimuth (straight ahead) or 45 degrees to the right. They presented two sequential vowels, which were either identical or differed in either spatial location or phonetic identity. They measured brain activity using fMRI and MEG when participants attended to spatial or phonetic attributes of the vowels. Attentional orientation significantly modulated the BOLD signal in auditory cortices. Regions specialised for spatial processing, such as posterior temporal cortex and posterior parietal regions, displayed significantly greater activity when attending to location; whereas attending to phoneme identity increased activity in anterior and superior temporal cortex. This pattern of results is consistent with the finding that specialised auditory pathways process spatial and identity information (Adriani et al., 2003; Ahveninen et al., 2006; Alain, Arnott, Hevenor, Graham, & Grady, 2001; Leavitt, Molholm, Gomez-Ramirez, & Foxe, 2011; Warren & Griffiths, 2003).

Attentional modulations of activity have also been observed at a single-unit level. For example, Lee and Middlebrooks (2011) recorded the spatial sensitivity of A1 neurons in cats when they performed two different tasks. One task was an active listening task in which they were required to detect stimulus periodicity, but spatial location was irrelevant. In the other task, they had to detect differences in location between two Gaussian noise bursts. A sub-set of A1 neurons (26–44%) showed sharpened spatial tuning in the location task compared to the periodicity task. This result demonstrates that the response selectivity of individual neurons can be influenced by spatial attention. Modulations of A1 spectro-temporal response properties have also been found when ferrets attend to a particular frequency (Elhilali, Fritz, Chi, & Shamma, 2007; Fritz, Shamma, Elhilali, & Klein, 2003).

1.1.3.3. Multi-talker listening

Multi-talker listening activates a similar fronto-parietal network as other types of acoustical stimuli (Hill & Miller, 2010). There has been recent interest in decoding the direction of attention from the neural response. For example, Ding and Simon (2012) presented speech from opposite-gender talkers to the left and right ears through headphones. They recorded activity using MEG and extracted spectro-temporal envelopes from source activity that was estimated to be consistent across trials. Envelopes reconstructed from low-frequency (< 8 Hz) activity resembled the attended speech more than the unattended speech on 92% of trials. This result demonstrates that the attended stimulus is preferentially encoded in MEG source activity.

The direction of attention can also be decoded from the neural response when there is no difference in the spatial location from which simultaneous speech is presented. For example, Mesgarani and Chang (2012) presented participants with two opposite-gender talkers, but monaurally from a single loudspeaker. Participants were asked to report key words spoken by a target talker. High-frequency gamma activity (75–150 Hz) was recorded from multi-electrode recordings placed over the dorsolateral temporal lobe in patients undergoing epilepsy surgery. A linear classifier was trained on examples of neural responses to individually-presented talkers. When the two talkers were presented simultaneously, the classifier was able to decode the identity of the target talker and the words that participants reported with 93% accuracy. On trials in which the participant responded incorrectly, the neural response did not always classify the target speech, but instead showed a tendency to identify the masking speech. These results shows that attended stimuli are preferentially encoded in the cortical response over unattended stimuli.

Together, the results of Ding and Simon (2012) and Mesgarani and Chang (2012) demonstrate that both high (75–150 Hz) and low-frequency (< 8 Hz) cortical responses reflect the direction of attention. It has been suggested that low-frequency responses might correspond to low-frequency modulations in the speech signal (Greenberg & Ainsworth, 2006; Rosen, 1992). Whereas, decoding attention from high-frequency responses is consistent with the finding of increased gamma-band activity in sensory regions when *visual* attention is directed inside a cell's receptive field (Fries, 2001; Gregoriou et al., 2009). Synchronised gamma activity has been suggested as a possible attentional mechanism by which activity in prefrontal regions affects sensory responses (see Section 1.1.3.2).

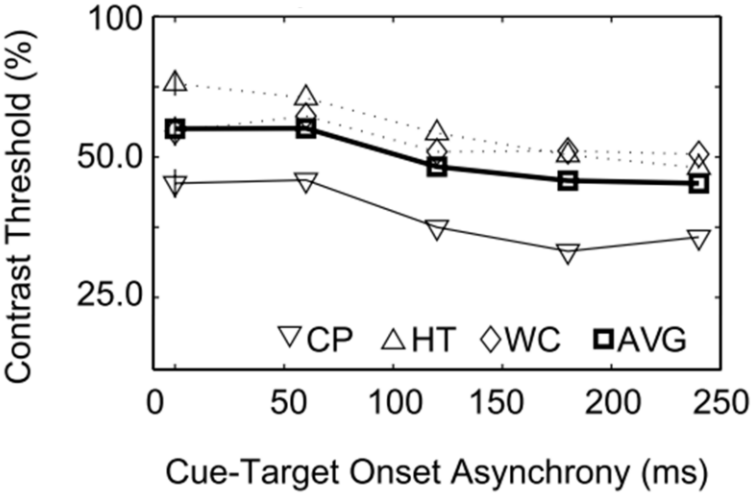


Figure 1.1. Adapted from Lu *et al.* (2009). Graph showing contrast thresholds when participants received a central visual cue at one of five cue-target intervals. Individual participant data are displayed with triangles and diamonds. Squares display averages across the three observers.

1.1.4. Preparatory attention

When participants are cued to attend to a particular dimension, task performance improves when they have more time to 'prepare' for the target stimulus. This research dates back to the use of 'warning signals' that cued exogenous attention. Woodrow (1914) asked participants to respond to a target sound as quickly as possible and varied the amount of time that a warning signal preceded the target, between 1 and 24 seconds. The shortest RTs occurred with a stimulus-onset asynchrony of 2 seconds. For longer durations, RTs lengthened with increasing duration. This result indicates that there may be an optimal interval by which preparation improves the accuracy of detection. Interestingly, recent experiments have used much shorter durations—typically less than 1 second. Within the preparatory interval, neuroimaging results have revealed preparatory brain activity before a target is presented. This section will discuss behavioural and neural results for preparatory endogenous attention in vision and audition.

1.1.4.1. Visual preparation benefits performance

When an endogenous cue is presented before a target visual stimulus, the cue-target interval affects RTs. Lu *et al.* (2009) presented a central white arrow to indicate the target location, prior to the presentation of four different Gabor patches at different spatial locations. The stimulus onset asynchrony between the cue and the target (i.e. the cue-target interval) was varied between 0 and 240 ms. Participants discriminated the orientation of target Gabor patches and their contrast thresholds were recorded. Figure 1.1 shows that contrast thresholds were better for longer cue-target intervals. A similar pattern of results was reported by Yamaguchi, Tsuchiya, and Kobayashi (1994) using a detection task and a set of longer cue-target intervals (200, 500, and 800 ms). A significant difference in RTs for detecting the target stimulus was found between all three cue-target intervals. Together, these studies suggest that participants gain a behavioural advantage from having more time to prepare for an upcoming visual stimulus, up to at least 800 ms.

1.1.4.2. Neural basis of visual preparation

During preparatory attention, a fronto-parietal network is active when participants are cued to spatial (location) and non-spatial (colour) attributes. This network overlaps with the regions reported during selective attention (during the target array). During the cue-target interval, some regions of the network are invoked by both location and colour cues. These regions may, therefore, contribute to domain-general attentional preparation. They include dorsal parietal cortex, dorsal posterior frontal cortex, and medial frontal cortex (Giesbrecht et al., 2003; Slagter et al., 2007; Woldorff et al., 2004). Cue-specific activity is also observed in dorsal and ventral cortical regions that are specialised for processing the cued dimension (Giesbrecht et al., 2003; Slagter et al., 2007; Woldorff et al., 2004). In these regions, the amplitude of pre-target BOLD activity correlates with behavioural performance (Giesbrecht, Weissman, Woldorff, & Mangun, 2006). Taken together, these studies suggest that similar brain regions are active during preparatory as during selective attention.

The time course of activity was explored by Green and McDonald (2008) using EEG. They presented a coloured cue on each trial that informed participants of the likely target location (80% validity). The target stimulus consisted of five bars aligned horizontally or vertically, which were displayed for 50 ms before being masked. Participants had to discriminate the orientation of the target stimulus. The researchers analysed EEG responses to the cue before the target was presented. During the first 300 ms after cue onset, the results showed feedforward activation that began in extrastriate cortex and moved upwards to the superior and inferior parietal lobes before reaching the frontal lobes. Following this initial activation, posterior parietal cortex became active for a second time and activity returned to extrastriate regions. The authors concluded that the visual cue sends feedforward activity to the fronto-parietal network, which then sends feedback activity to extrastriate regions. They found that the magnitude of activity within regions of the fronto-parietal network strongly predicted the accuracy with which participants discriminated target orientation ($r \geq 0.78$). Together with the results of Giesbrecht et al. (2006), this result demonstrates that pre-target activity, both in visual and fronto-parietal regions, predicts task performance.

1.1.4.3. Auditory preparation benefits performance

Preparation has also been found to improve the accuracy of pitch discrimination for an acoustical target stimulus. Arbogast and Kidd (2000) presented pure tones from seven loudspeakers arranged in a semi-circle around the participant. On each trial, one tone was

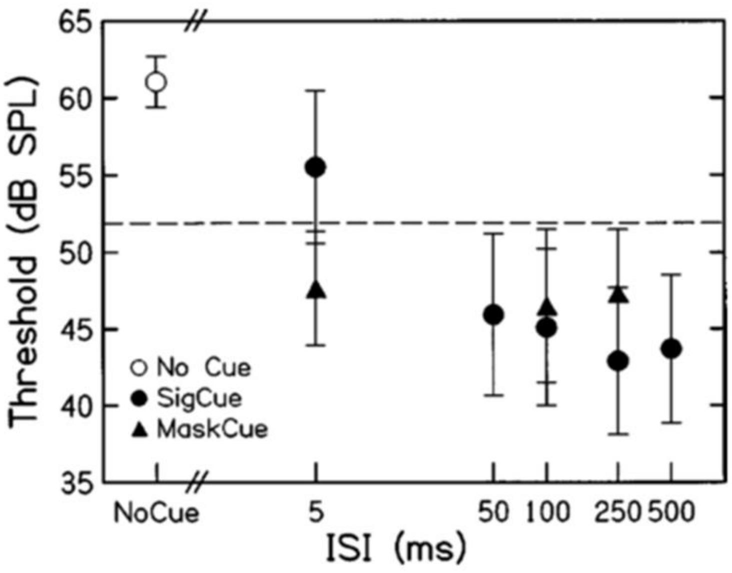


Figure 1.2. From Richards and Neff (2004). Graph showing average thresholds for the no-cue condition, the signal-cue condition, and the masker-cue condition. For the latter two conditions, thresholds are plotted as a function of ISI. Error bars show the standard error of the mean across 6 observers. The dashed line shows the signal level required to increase the masker stimulus by 1 dB.

presented from a target loudspeaker and different tones were presented simultaneously from masking locations. Participants had to detect whether the target had a rising or falling pitch. Participants were cued to attend to one loudspeaker location for each block of trials. When the target was presented from the cued loudspeaker (75% of trials), responses were more accurate and had shorter latencies than when the target was presented from an uncued loudspeaker.

In the auditory modality, it is unclear how the duration of the cue-target interval affects performance. One study by Richards and Neff (2004) systematically varied the amount of time between the offset of a cue and the onset of the target array (inter-stimulus interval: ISI). They measured thresholds for detecting a 1000 Hz pure-tone target in the presence of multi-tone complex maskers. There were three cueing conditions. The participant either received a preview of the signal ('signal-cue'), a preview of the masker ('masker-cue'), or no cue. The ISI was varied between 5 ms and 500 ms. When participants received either the signal-cue or the masker-cue, thresholds were better than in the no cue condition. For the signal-cue condition, detection thresholds were worse for the 5 ms interval than all of the other intervals (Figure 1.2). This result suggests that participants gain some benefit from having 50 ms to prepare for the target over 5 ms, but receive no further improvement between 50 and 500 ms. When participants received the masker-cue, there was no effect of increasing the ISI. However, since only short (≤ 500 ms) ISIs were used in this experiment, it is not clear whether longer preparation times would improve detection thresholds further.

1.1.4.4. Neural basis of auditory preparation

Similar to vision, preparing for an acoustical stimulus increases pre-target activity in a fronto-parietal network. Wu, Weissman, Roberts, and Woldorff (2007) presented an acoustical cue that instructed participants to attend to the left or right. Participants had to detect a monaural tone pip occurring on that side. The cue evoked activity in auditory cortex and in a medial-superior fronto-parietal network, which included frontal gyri, anterior cingulate cortex, and the superior parietal lobule. The researchers comment that the distribution of the fronto-parietal network was slightly more superior and medial than was reported in a similar visual study (Woldorff et al., 2004). However, a within-experiment manipulation is required to confirm this observation. In general, many of the fronto-parietal regions overlapped with those observed during visual preparation.

Voisin, Bidet-Caulet, Bertrand, and Fonlupt (2006) also showed modulation of activity in auditory cortex. An arrow cued attention to the left or to the right and participants had to detect the presence of a noise burst that emerged with increasing intensity. Results showed BOLD activity in regions of the fronto-parietal network, including areas of the right dorsolateral prefrontal cortex and inferior frontal cortex, when participants were cued to the left or the right. However, contrasts between left and right trials revealed activity in the superior temporal sulcus (including Heschl's gyrus and surrounding areas) that occurred contralateral to the cued side. This study, therefore, provides evidence for cue-specific preparation in auditory cortex.

1.1.5. Summary and conclusions

This section has highlighted the benefits of directing attention to acoustical and visual stimuli for RTs and, in some instances, accuracy. Converging evidence shows activation of a fronto-parietal network during selective attention, which controls the allocation of attentional resources to regions specialised for processing the relevant dimension. This network is active in auditory and visual modalities and contributes to domain-general attention.

In both modalities, the latency with which stimuli can be accurately discriminated depends on the duration of time provided for participants to prepare for cued attributes of the target stimulus before it is presented. However, the time course of 'preparatory attention' is not fully understood. In particular, it is unclear whether longer cue-target intervals (> 800 ms) improve performance further.

When a visual cue is presented before the target stimulus, activity occurs first in extrastriate cortex and next in the fronto-parietal network, which then feeds back to sensory regions that are specialised for processing the relevant dimension (Green & McDonald, 2008).

Activity within sensory regions is cue-specific and depends on the attended attribute. Activity in fronto-parietal and cue-specific sensory regions *both* correlate with the accuracy of visual orientation discrimination (Giesbrecht et al., 2006; Green & McDonald, 2008), suggesting that activity in both of these regions are beneficial for successful attentional preparation.

1.2. Attention during multi-talker listening

In everyday situations, listeners often face the challenge of identifying speech against a background of competing voices (Darwin, 2008). The ability to attend selectively to speech enables communication in these settings (Shinn-Cunningham, 2008). This section will first discuss how different types of background noise and characteristics of multi-talker listening affect speech intelligibility. Second, this section will discuss the improvement in the accuracy of speech intelligibility gained from 'preparatory attention', when participants know characteristics of the target talker before he or she begins to speak.

1.2.1. Different types of masker

Background noise negatively impacts speech intelligibility—both for speech masked by random noise (Festen & Plomp, 1990; Song, Skoe, Banai, & Kraus, 2011) and for speech masked by other speech (Festen & Plomp, 1990). Speech intelligibility depends on a complex relationship between the number of competing sources, the spectro-temporal properties of the sources, and spatial configuration of sources (see Bronkhorst, 2000 for a review).

1.2.1.1. Energetic and informational masking

Energetic masking is determined by the relationship of the frequency spectrums between the noise and the signal and their relative intensity (Scott, Rosen, Wickham, & Wise, 2004). It is thought to arise due to competition at the auditory periphery (Scott, Rosen, Beaman, Davis, & Wise, 2009). Energetic masking explains the well-known finding that the intelligibility of a talker masked by flat-spectrum noise decreases as the target-to-masker SNR decreases. Darwin (2006) proposed that energetic masking is a problem of detecting the spectrum of the target above the masker in relevant frequency bands. The detrimental effects of energetic masking are thought to arise at the stage of auditory processing where frequency analysis is performed on the basilar membrane of the cochlea (Scott et al., 2004).

Energetic masking predicts that continuous noise is a more effective masker of target speech than a single competing talker, but there are also conditions in which competing speech produces additional masking over noise maskers (Brungart, Simpson, Ericson, & Scott, 2001; Freyman, Helfer, McCall, & Clifton, 1999). Masking that is not due to energetic masking has been termed 'informational' masking (Brungart & Simpson, 2002a; Freyman, Balakrishnan, & Helfer, 2004). Informational masking is thought to result from confusion about which spectro-temporal information belongs to the target and masker (Durlach et al., 2003). Scott *et al.* (2009) proposed that informational masking is a result of competition in central auditory processing. This idea is consistent with errors for speech-in-speech tasks. On incorrect trials, participants frequently report words from the masking talker, rather than random words from the response set, even when the masker is presented at the same intensity as the target (Brungart & Simpson, 2002a; Darwin, 2006). This finding suggests that the accuracy of speech intelligibility can be impaired by confusion between the target and masker. This characterisation of informational masking predicts that only maskers consisting of intelligible speech sounds will show informational masking and, furthermore, that the amount of informational masking should not be affected by SNR. Several studies show that the intelligibility of attended speech masked by other speech is relatively constant over a range of SNRs from -12 to 0 dB, and only increases when the SNR is increasingly positive (Brungart, 2001; Makeig, Debener, Onton, & Delorme, 2004). Of course, a speech masker provides both energetic and informational masking, which is why SNR still influences intelligibility at positive ratios.

1.2.1.2. Single-talker masker

The amount of energetic masking provided by a single-talker masker depends on the frequency spectrums of the target and masking speech. Male and female talkers have different average fundamental frequencies (F0s) and vocal-tract lengths, which contribute to differences in the frequency spectra of speech produced by male and female talkers (Smith & Patterson, 2005). Normally-hearing listeners are surprisingly good at segregating two talkers (Allen, Alais, Shinn-Cunningham, & Carlile, 2011), even when the competing talker has a similar spectral profile as the target or is presented at a higher intensity (Brungart, 2001).

One factor that might make single-talker speech a poor masker is the inherent fluctuations that are present—including fluctuations in amplitude and spectrum (Freyman et al., 2004). Listeners take advantage of the portions of a target signal where the level of the masker is low and the SNR is favourable, often referred to as ‘dips’ or ‘glimpses’ (Cooke, 2006; Moore, 2008b). For normally-hearing listeners, the proportion of target glimpses in modulated noise is a good predictor of intelligibility (Cooke, 2006). This result demonstrates that the temporal distribution of energy in the masker is important for determining how much masking occurs.

1.2.1.3. Multi-talker masker

Multiple-talker maskers typically result in lower accuracy of speech intelligibility than single-talker maskers. Glimpses are not correlated across speech signals, so increasing the number of competing talkers decreases the portion of the target that contains glimpses. Increasing the number of talkers produces a spectro-temporal profile that increasingly resembles steady-state noise (Bronkhorst, 2000), where there are few glimpses with four or more talkers (Bronkhorst & Plomp, 1992). At the same SNR, the accuracy with which target words are identified decreases as the number of interfering talkers increases (Nelson, Bolia, Ericson, & Mckinley, 1998; Yost, Dye, & Sheft, 1996). This finding is consistent with the fact that maskers containing more talkers contain progressively fewer glimpses and the finding that the proportion of glimpses in modulated noise predicts intelligibility (Cooke, 2006).

1.2.2. Cues for segregation

1.2.2.1. Which types of cues can listeners utilise?

The difficulty of multi-talker listening depends not only on the type of masker (single-talker or multiple-talker), but also on the perceptual similarity of the target and masker(s). Multi-talker listening is most difficult when the target and maskers are perceptually similar. These perceptual attributes include voice characteristics (which encompasses a wide range of cues, such as fundamental frequency, vocal tract size, accent, speaking style, timbre and amplitude modulation), perceived spatial location, intensity, timing, and lexical content. Separating the target and masker on any of these dimensions improves intelligibility (Allen, Carlile, & Alais, 2008; Brungart et al., 2001; Münte et al., 2010; Varghese, Ozmeral, Best, & Shinn-Cunningham, 2012; Xiang, Simon, & Elhilali, 2010).

1.2.2.2. How do cues benefit listeners?

Ensuring that targets and maskers are perceptually distinct can offer several advantages. First, if perceptual cues reduce the overlap of spectro-temporal features, then energetic masking will be reduced (Brungart & Simpson, 2002a). Secondly, perceptual cues may improve a listener’s ability to direct selective attention to the target talker (Shinn-Cunningham, 2008; Varghese et al., 2012), which may reduce informational masking (Freyman et al., 2004; Shinn-Cunningham, 2005).

1.2.2.3. Differences in location

Introducing a difference in location between targets and maskers has been found consistently to improve the intelligibility of a target talker (Bronkhorst & Plomp, 1988; Darwin & Hukin, 1999; Helfer & Freyman, 2005). Spatial localisation is mediated by three main cues: (1) differences in the *timing* of waveforms between the two ears, known as inter-aural timing differences (ITDs); (2) differences in the *level* of waveforms between the two ears, known as inter-aural level differences (ILDs); and (3) monaural spectral cues.

Participants are able to utilise differences in the perceived location of sounds to segregate talkers during multi-talker listening.

The term 'spatial release from masking' (SRM) describes the decrease in SNR that can be applied to a spatially separated target and masker to produce the same behavioural performance as when the sounds are collocated. Allen et al. (2008) asked participants to report key words spoken by a target talker in the presence of two masking talkers, both when the talkers were collocated and when they were spatially separated. They reported a SRM of 12 dB when the maskers were located 30° to the left and right of the target. This finding means that the SNR could be 12 dB lower in the spatially separated condition, while producing accuracy that was equivalent to the collocated condition.

In particular, Allen et al. (2008) suggested that spatial separation in the azimuthal (horizontal) plane may be important for the initial segregation of talkers. They found that even when two voices were initially spatially-separated and gradually became collocated, participants showed 4 dB release from masking. Importantly, the key words were spoken after the voices became collocated. This result shows that a difference in location helps listeners to segregate the target talker from the interfering talker and, furthermore, that listeners can use an initial location difference to segregate simultaneous talkers even when that location difference disappears. This strategy is plausible in real-life situations, where listeners often need to track moving talkers. Importantly, Allen *et al.* found that release from masking was greater for opposite-gender than same-gender talkers who were initially spatially separated. This result suggests that initial spatial separation may allow additional benefit to be gained from targets and maskers that are also separated in F0, although initial spatial separation may not be able to be utilised in the absence of other cues.

The *degree* to which simultaneous talkers are spatially separated also affects speech intelligibility (Marrone, Mason, & Kidd, 2008c; Noble & Perrett, 2002). For example, Noble and Perrett (2002) presented participants with continuous speech from 0° azimuth (straight ahead) in the presence of two distracting talkers. Intelligibility was better when the distracters were located at ± 30° from the target than when they were collocated. However, intelligibility improved further when the distracters were located at ± 90° (Figure 1.3). This result demonstrates that the degree of spatial separation is important in determining release from masking.

1.2.2.4. Differences in fundamental frequency

A difference in F0 also improves the ease with which competing talkers can be segregated. F0 varies considerably, both within and between talkers. Along with vocal tract length, F0 provides evidence for a talker's gender and age (Bachorowski & Owren, 1999; Iseli,

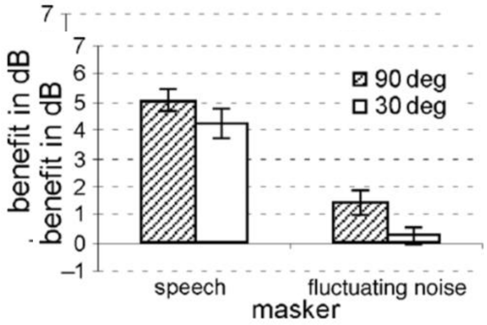


Figure 1.3. From Noble and Perrett (2002). Benefit (in dB) of spatially separating the target speech from a competing sound (speech or fluctuating noise) at two levels of spatial separation (± 30° and ± 90°). Positive values indicate better performance in the spatially-separated condition.

Shue, & Alwan, 2007; Murry & Singh, 1980). Females have higher F0s than males (Bachorowski & Owren, 1999), shorter vocal tract lengths (Bachorowski & Owren, 1999), and their formant frequencies are, on average, 16% higher (Peterson & Barney, 1952).

Speech intelligibility during two-talker listening is more accurate for opposite-gender talkers than for same-gender talkers (Figure 1.4; Brungart et al., 2001; Brungart, 2001; Shafiro & Gygi, 2007). Also, participants subjectively rate same-gender talkers as more difficult to segregate than opposite-gender talkers (Nakai, Kato, & Matsuo, 2005). When participants are asked to identify words spoken by a target talker, nearly all incorrect responses consist of words spoken by the competing talker (Darwin, Brungart, & Simpson, 2003). This finding suggests that the primary benefit of presenting opposite-gender talkers, rather than same-gender talkers, is release from informational masking.

Even for same-gender talkers, speech intelligibility improves with larger F0 differences (Darwin & Hukin, 2000). Although, previous findings demonstrate that participants can utilise even small F0 differences to segregate talkers. For example, Assmann and Summerfield (1994) presented participants with brief 50 ms synthetic vowels that

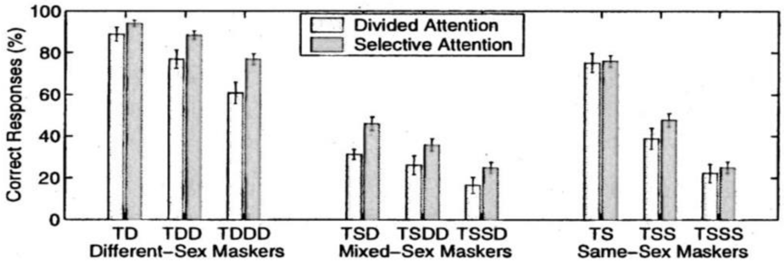


Figure 1.4. From Brungart *et al.* (2001). Graph shows percentage of correct identification of key words spoken by a target talker in three masking conditions. The different-sex masker condition contained either 1 (TD), 2 (TDD), or 3 (TDDD) different-sex maskers. The mixed-sex masker condition contained 1 same- and 1 different-sex masker (TSD), or an additional different-sex (TSDD) or same-sex (TSSD) masker. The same-sex masking condition contained either 1 (TS), 2 (TSS), or 3 (TSSS) same-sex maskers. Error bars show 95% confidence intervals in each condition.

differed in F0 and asked them to identify the presented vowels. Listeners gained a greater benefit for vowel intelligibility from differences in F0 when the stimulus had a longer duration compared to a short duration. Also, when the difference in F0 was small (0.25-1 semitone), presenting at least two different segments of the vowel produced higher accuracy than presenting the same segment repeatedly. This result suggests that the limit on intelligibility for short-duration stimuli is not the time taken to switch attention between the two segments. Moreover, intelligibility depended on the particular time segment of the vowel that was presented. The researchers suggested that, for small differences in F0, listeners perform a sequence of analyses of different time-segments of a vowel to determine where the formants are most clearly defined.

1.2.2.5. Interaction between cue types

The previous sections have highlighted the advantages that can be gained from separating concurrent talkers in *either* spatial location or F0. However, these factors can also be used in combination. Shomstein and Yantis (2006) presented two talkers through headphones. They observed higher intelligibility if a talker and opposite-gender distracter were also presented to separate ears, rather than binaurally with the same perceived location. This result is consistent with the idea that listeners track a combination of different cue types (Mondor, Zatorre, & Terrio, 1998).

The finding that listeners benefit from access to multiple cues is consistent with the finding of enhanced sensitivity to features of a source that are not task-relevant. For example, Maddox and Shinn-Cunningham (2012) presented two simultaneous digit streams, which were preceded by a ‘primer’ phrase. Participants were instructed to report the digit stream that matched the primer phrase in either location (left/right) or pitch (high/low) and to ignore the task-irrelevant feature. On each trial, one digit of each simultaneous pair matched the location of the primer phrase and one digit matched its location. On consistent trials, the digits at the location of the primer phrase shared the pitch of the primer. On mix trials, the task-irrelevant feature varied within each digit sequence. Overall, participants were more likely to correctly report the digit stream on consistent trials than mix trials. This finding demonstrates that a task-irrelevant feature can influence the accuracy of speech intelligibility. This result is consistent with object-based attention, which suggests that participants attend to spatial and non-spatial attributes of stimuli in combination (see Section 1.1.2.2).

When identifying words spoken during multi-talker listening in everyday life, it might be advantageous to monitor multiple cues at once rather than focusing on a single cue. The dynamic nature of speech signals means that the factors most useful for segregating talkers might vary over time (Caporello Bluvas & Gentner, 2013). For example, F0 fluctuations in the speech spoken by one talker or changes in location for a talker who moves while speaking are both factors that might cause the most useful cue for talker segregation to change over time. Tracking both location and F0 cues may also be important when a target talker cannot be distinguished from multiple distracting talkers by either information source alone, but only by combined knowledge of location and F0.

1.2.3. Preparatory attention during multi-talker listening

Normally-hearing listeners are able to utilise information about a target talker before he or she begins to speak to improve the intelligibility of that talker. Advantages have been demonstrated from knowing the spatial location (Best et al., 2009; Best, Ozmeral, & Shinn-Cunningham, 2007; Ericson, Brungart, & Brian, 2004; Kidd, Arbogast, Mason, & Gallun, 2005) and the identity and timing (Kitterick et al., 2010) of the target talker. In these experiments, the stimulus composition is identical on every trial, but the participant is instructed to attend to different talkers on different trials.

1.2.3.1. Advantages of preparing for multi-talker listening

Several experiments demonstrate improved intelligibility in multi-talker listening when participants know a talker's spatial location before he or she begins to speak (Best et al., 2009; Best, Ozmeral, et al., 2007). For example, Best *et al.* (2007) presented participants with a sequence of spoken digits and five maskers containing time-reversed speech. The target and maskers were presented simultaneously from five different loudspeakers and participants had to report the target digits in the correct order. On cued trials, a light-emitting diode was presented above one of the loudspeakers to inform participants of the location of the upcoming digits. Identification accuracy was significantly higher on cued trials than uncued trials.

Exposing participants to characteristics of the upcoming target voice also improves intelligibility (Freyman et al., 2004; Kitterick et al., 2010). For example, Brungart *et al.* (2001) found that prior experience of the target talker significantly reduced the number of opposite-gender confusions compared to when participants did not have prior experience of the target talker. This result suggests that knowing characteristics of a target talker provides a release from informational masking. In this experiment, the greatest benefit of prior experience for the accuracy of speech intelligibility occurred in the three- and four-talker listening conditions. In the two-talker condition, accuracy was near-ceiling even without the cue.

Therefore, prior experience of a talker may be most beneficial in difficult listening situations when intelligibility is low.

Together, the results discussed in this section show that prior knowledge of a target talker's location or exposure to their voice can improve intelligibility during multi-talker listening compared to when participants do not know this information in advance. Allen *et al.* (2011) also propose that intelligibility is affected by prior knowledge of the *masker* location. They found that participants were worse at identifying phonemes spoken by a target talker when masking phonemes were presented from unexpected locations than expected ones. On average, thresholds for identifying phonemes were 2.6 dB higher when maskers were presented from unexpected locations.

1.2.3.2. Timing of preparatory attention during multi-talker listening

Although previous research has consistently demonstrated that knowing characteristics of an upcoming talker improves intelligibility, little is known about the time course of preparatory attention. It is possible that participants wait until just before the target begins to speak to prepare their attention (Liu, 2003). Thus, allowing more time to prepare would not improve intelligibility. Alternatively, preparatory attention may unfold over time, such that increasing the amount of preparation time progressively improves intelligibility. Previous experiments have tended to use different cue-target intervals, ranging between 100 ms before the target (Koch, Lawo, Fels, & Vorländer, 2011) to cueing at the beginning of each block (Brungart & Simpson, 2007; Ericson et al., 2004; Kitterick et al., 2010). No similar experiments have systematically varied the cue-target interval within a single experiment.

Different time intervals have been investigated in experiments in which participants are asked to switch their attention from one attribute to another. It is well-established that there is a switch cost—that is, RTs are longer when participants have to switch attention to a different attribute than when participants maintain attention on the same attribute (S Monsell & Driver, 2000; Rogers & Monsell, 1995). Several experiments have demonstrated that the switch cost is reduced when participants are given longer intervals over which to switch their attention (Koch et al., 2011; Larson & Lee, 2013a; Meiran, Chorev, & Sapir, 2000). For example, Larson and Lee (2013a) presented participants with two simultaneous sequences of spoken letters. The sequences were monotonised and shifted in pitch to produce an 8.5 semitone difference. After the first three digits were spoken, there was a silent gap where participants either had to switch to the opposite sequence or maintain their attention on the same sequence. Participants' task was to detect the second "E" that occurred in the attended sequence. Before the letters began, participants received an auditory preview

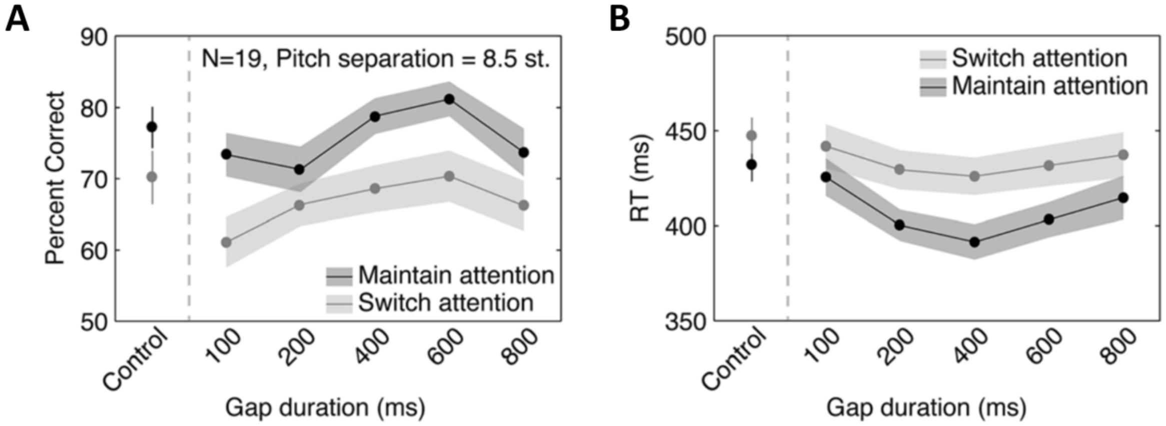


Figure 1.5. From Larson and Lee (2013). Graph shows accuracy (**A**) and RTs (**B**), as a function of the gap duration allowed for switching attention. Shaded areas show ±1 standard error. The "Control" condition pooled trials where the second target "E" occurred within the first three letters, i.e., before the gap.

of the initial target talker and a cue that told them whether to stick or switch in the gap. There were five different gap durations: 100, 200, 400, 600 and 800 ms. The results showed a significant main effect of gap duration on accuracy. Accuracy was significantly lower for the shorter durations (100 and 200 ms) than the longer durations (400 and 600 ms), but decreased when the duration was increased further to 800 ms (Figure 1.5A, "Switch attention" condition). A similar pattern of results was found for RTs (Figure 1.5B). This result shows that increasing the duration of a gap improves accuracy and latency, although intervals greater than 400 ms did not lead to increased accuracy or latency.

The switch cost itself has sometimes been thought to reflect preparatory processes (Meiran et al., 2000). Therefore, it is possible that similar processes underlie attentional preparation for an upcoming talker as switching attention within a gap. The key difference between the experiment of Larson and Lee (2013a) and experiments that provide an instructional cue for which attention is sustained throughout the trial, is that Larson and Lee varied the duration of time provided to switch attention from one attribute to another (but the interval between the presentation of the cue and the onset of the gap remained constant), rather than the amount of time between the cues and the onset of the talkers. Nevertheless, if similar processes occur during both intervals, the results of Larson and Lee suggest that longer durations of preparation time during the cue-target interval have the potential to improve the accuracy of speech intelligibility until an 'optimal' time, after which longer intervals decrease accuracy.

1.2.4. Preparatory brain activity for multi-talker listening

Preparing for an upcoming talker in a mixture of talkers invokes brain activity in a fronto-parietal network similar to that observed for visual and acoustical non-speech stimuli (Hill & Miller, 2010; Lee et al., 2013). Hill and Miller (2010) measured brain activity using fMRI. On each trial, three simultaneous talkers were presented, which differed in simulated spatial location and average F0. Before the acoustical stimuli began, a visual cue indicated either the location (left/right/centre) or the F0 (high/low/middle) of the target talker. Participants' task was to press a button when they detected a sentence onset in the attended speech stream. Visual cues for both location and gender evoked activity in the same left-hemisphere fronto-parietal network. However, the detailed pattern of activity within the network depended on whether participants were preparing to select the upcoming talker by location or F0. The dorsal precentral sulcus and superior parietal lobule displayed significantly greater BOLD activation when participants were cued to location, whereas the inferior frontal gyrus displayed significantly greater activation when participants were cued to pitch (Figure 1.6). Thus, the results provide evidence for both domain-general and cue-specific brain activity, as has been observed for visual and acoustical non-speech stimuli.

A similar experiment by Lee *et al.* (2013) measured preparatory activity using magneto-encephalography (MEG). On each trial, two digits were spoken simultaneously, which differed in simulated spatial location (left/right) and F0 (high/low). A visual cue, indicating either the spatial location or F0 of the target talker, preceded the acoustical stimuli. Lee *et al.* found greater activity in the left dorsal precentral sulcus and gyrus during attend-location trials and greater activity in the left posterior superior temporal sulcus during attend-F0 trials (Figure 1.7). These results are compatible with Hill and Miller's, which show modulations in similar brain regions. The results are also compatible with the role of the superior temporal sulcus in voice identification (Belin & Zatorre, 2003).

A possible shortcoming of the experiments of Hill and Miller (2010) and Lee *et al.* (2013) is that differences in the feature to be used for selection (e.g. location or F0) were confounded with differences in the visual cues. Both studies used horizontal arrows to cue location and vertical arrows to cue F0. Woldorff *et al.* (2004) argue that cue-triggered activity could *either* arise from attentional processing of the cues or from sensory and semantic processing necessary to interpret the cues. In the experiments of Hill and Miller and Lee *et al.*, no control condition was implemented to establish whether the physical difference in cue orientation contributed to the observed differences in brain activity.

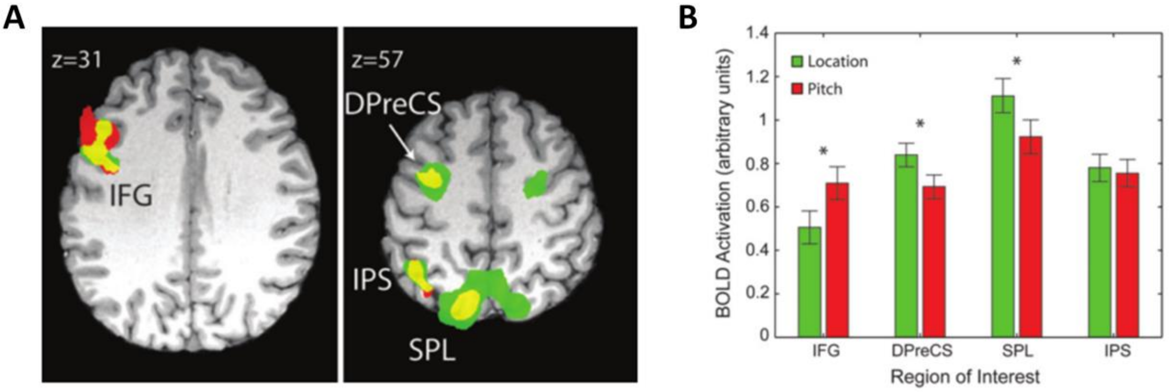


Figure 1.6. Adapted from Hill and Miller (2010). (**A**) BOLD activity in response to a visual cue for attention (preparatory attention). Figures show attend-F0 (red) and attend-location (green) conditions, each contrasted with a rest cue condition, and the overlap between attend-F0 and attend-location conditions (yellow). Activations include inferior frontal gyrus (IFG), dorsal precentral sulcus (DPreCS), inferior parietal sulcus (IPS), and superior parietal lobule (SPL; all $p < 0.01$ FDR). Activations are shown on the brain of a representative participant from this experiment. Graph **B** shows a region of interest (ROI) analysis. The graph illustrates BOLD activity in attend-F0 and attend-location conditions. Error bars show the standard error of the difference within subjects. ROIs were defined by overlap between F0 and location cues greater than rest cues (the yellow regions in **A**). All tests are two-tailed paired *t*-tests (*$p < 0.05$, uncorrected).

Neither Hill and Miller (2010) nor Lee *et al.* (2013) addressed the question of how early in time attentional preparation is manifest in neural activity. Hill and Miller's experiment revealed brain activity only with the low temporal resolution of fMRI. Lee *et al.* did not analyse MEG data until 600 ms after the start of the visual cues. (They displayed the visual cue together with a fixation dot for 300 ms; they then extinguished the cue, leaving only the dot for 700 ms, at which point the acoustical stimuli were presented. They analysed MEG data in 400-ms windows immediately before and after the onset of the acoustical stimuli. Thus, 600 ms elapsed between the onset of the visual cue and the start of the first analysis window.) An experiment investigating preparation for an upcoming *visual* stimulus revealed brain activity less than 250 ms after the cue was presented (Yamaguchi et al., 1994). This result from vision suggests that attention has the potential to influence preparatory brain activity with latencies shorter than 600 ms.

Srinivasan, Thorpe, Deng, Lappas, and D'Zmura (2009) showed that the spectral features of the EEG signal can be used to decode the direction of attention during the cue-target interval. They presented two male talkers simultaneously on each trial. An acoustical cue that preceded the talkers instructed participants to attend to the left (-45°) or right (+45°) talker. They found optimal classification accuracy 400–900 ms after the onset of the

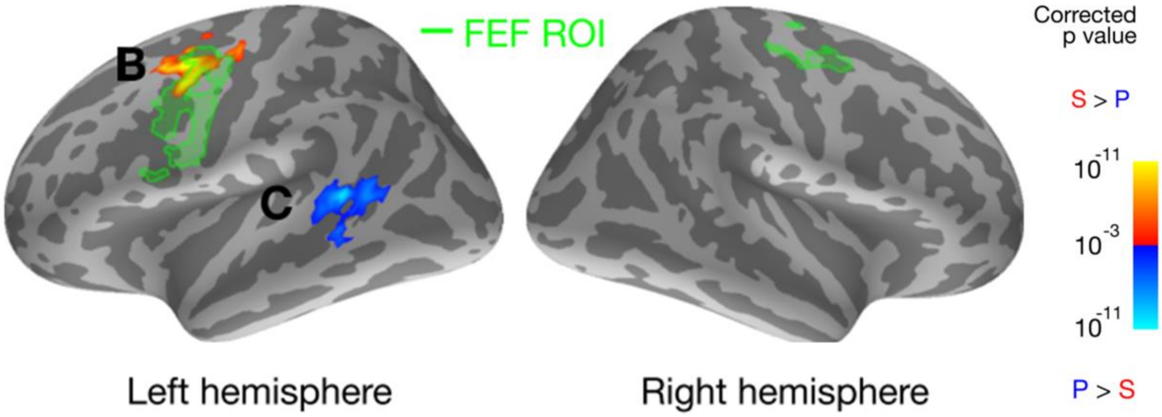


Figure 1.7. Adapted from Lee *et al.* (2013). Inflated surfaces of the left and right hemispheres show statistical maps illustrating a vertex-by-vertex comparison (with minimum cluster-size threshold set to 100 vertices) in response to the visual cue (preparatory attention). Red/yellow areas indicate greater activation in spatial location (S) than pitch (P) trials, while blue/turquoise areas show the opposite contrast. Surfaces show activation in left frontal eye fields (**B**) and left superior temporal gyrus (**C**). In green, is displayed the frontal eye field region of interest (FEF ROI) obtained from a memory-guided go/no-go saccade task.

cue (approximately 75% classification accuracy). This was the latest interval that they analysed. The results demonstrate that attentional orientation can be decoded even before a target talker begins to speak. The results also suggest that the orientation of attention is best represented in the EEG response from 400 ms after the onset of an attentional cue. This finding is consistent with the idea that participants utilise the cue-target interval to prepare their attention and further suggests that attention influences preparatory brain activity with latencies shorter than 600 ms.

1.2.5. Summary and conclusions

Speech intelligibility depends on a complex interplay between the type of masker(s), number of maskers, similarity of the target and masker(s) in spatial location and F0, and whether participants know characteristics of the target talker in advance. There is evidence that a fronto-parietal network controls the allocation of attentional resources in preparation for a target talker, similar to results for upcoming visual and acoustical non-speech stimuli (see Section 1.1.4.).

The mechanisms that underlie preparatory attention remain unclear. One key question is how much preparation time leads to the highest accuracy of speech intelligibility. There are two alternatives: (1) the time interval does not influence intelligibility until it reaches a certain threshold for successful preparation, beyond which increasing the time interval does not improve intelligibility further; or (2) increasing the amount of preparation time improves intelligibility progressively, at least below a certain ‘optimal’ interval. The latter hypothesis is consistent with the idea that participants can partially prepare using shorter intervals, but they are able to prepare more successfully when given more time. Experiments that have provided a cue in advance of the target talker consistently find that advance cueing benefits the accuracy of speech intelligibility during multi-talker listening, but the amount of preparation time has not been varied within a single experiment.

Similarly, the time course of brain activity during preparation is not fully understood. Of the two experiments that have previously studied preparatory brain activity for multi-talker listening, one experiment measured brain activity with the low temporal resolution of fMRI (Hill & Miller, 2010) and the other did not analyse activity until 600 ms after the cue was presented (Lee et al., 2013). If the timing of an advance cue affects the intelligibility of a target talker, then the timing of brain activity during preparation is likely to be important for understanding this improvement in speech intelligibility.

1.3. Multi-talker listening and hearing loss

Converging evidence from accuracy scores and self-report suggests that multi-talker listening is particularly challenging for listeners with sensorineural hearing loss (e.g. Dubno et al., 1984; Helfer & Freyman, 2008). This section discusses possible consequences of this difficulty and peripheral factors that might contribute. I also discuss the reduced speech intelligibility benefit of cues for location or F0 in hearing-impaired listeners.

1.3.1. Particular difficulty in noisy environments

Listeners with sensorineural hearing loss often complain that they find it difficult to understand speech in the presence of background noise (Dubno et al., 1984; Marrone, Mason, & Kidd, 2008b; Salvi et al., 2002). The background noise that patients typically refer to is other speech (Marrone et al., 2008b). The problem is not due to detection, since patients are able to detect that a person is speaking to them, but they say that they are unable to understand what that person is saying (Salvi et al., 2002).

Accuracy scores for speech intelligibility reflect these complaints. When intelligibility is measured in terms of percent correct, listeners with sensorineural hearing loss perform much worse than normally-hearing listeners when identifying speech masked by competing speech (Helfer & Freyman, 2008). Furthermore, the difference in intelligibility between listeners with normal hearing and moderate hearing loss is greater for multi-talker listening than for speech recognition in quiet (Marrone et al., 2008a).

Difficulties with multi-talker listening can lead to social isolation in hearing-impaired listeners, since communication difficulties can lead to embarrassment and frustration in social settings (Shinn-Cunningham, 2007). Gatehouse and Noble (2004) administered the 'Speech, Spatial, and Qualities of Hearing Scale' (SSQ) to 153 patients with moderate hearing loss, who were seeking rehabilitation from an audiology clinic, prior to acoustic hearing aid fitting. They also administered a twelve-item handicap questionnaire, which provided a measure of social withdrawal and discomfort. Responses showed that one of the greatest perceived difficulties was background noise—specifically, simultaneous speech streams and listening in groups and noise. Patients also reported that perceiving speech in these situations was effortful. Handicap score significantly correlated with: following conversation with multiple people talking, ignoring interfering voices, and talking in noise. Therefore, difficulty encountered during multi-talker listening contributes strongly to perceived handicap in hearing-impaired listeners.

1.3.1.1. Hearing loss in older adults

The majority of research that has been conducted with hearing-impaired listeners has compared speech intelligibility for older adults (> 60 years) with hearing loss to young, normally-hearing adults. One problem with this comparison is that ageing is associated with a number of problems, including a decline in executive control (Braver & West, 2008; Huppert & Wilcock, 1997; Li, Lindenberger, & Sikström, 2001), which may be independent from, but correlated with, peripheral hearing loss. Declines in executive control or working memory, reduced speed of processing, or peripheral hearing loss all have the potential to impair performance on multi-talker listening tasks (Helfer & Freyman, 2008; Salthouse, 1996; Shinn-Cunningham & Best, 2008). Therefore, it is difficult to tease apart factors that result from peripheral hearing loss and cognitive decline.

Some researchers have suggested that peripheral hearing loss has associated central consequences, which might be the cause of cognitive decline in older adults (Pichora-Fuller, Schneider, & Daneman, 1995; Schneider, Daneman, & Murphy, 2005). Nevertheless, older adults also perform poorly on *visual* cognitive tasks (Kramer & Madden, 2008), suggesting that hearing loss is not the only reason for cognitive decline. The consequences of hearing loss and general cognitive decline in older adults might be separate factors that contribute to accuracy on speech intelligibility tasks. This idea is consistent with research showing a relationship between age and the accuracy of speech intelligibility in noise, even when pure-tone audiometry shows thresholds in the normally-hearing range (Dubno et al., 1984). Therefore, the reduced speech intelligibility of older adults is not necessarily due to reduced acoustic thresholds for detecting sounds.

It has also been reported that older adults show different neural responses to speech stimuli compared to younger adults. For example, Harkrider, Plyler, and Hedrick (2006) measured EEG activity during presentation of consonant-vowel stimuli in young normally-hearing adults (22–34 years), older normally-hearing adults (43–73 years), and older adults with moderate sensorineural hearing loss. The amplitude of the N1 component was significantly greater for both of the older adult groups than for young normally-hearing listeners. However, the difference in N1 amplitude between older normally-hearing adults and older hearing-impaired adults was not significant. This result demonstrates that there are differences in the neural response to speech between older and younger adults. Differences in neural response might be one factor that underlies poorer speech intelligibility in older adults, independent from increased audiometric thresholds that are associated with cochlear hearing loss. This idea highlights the fact that differences in the accuracy of speech intelligibility between young normally-hearing adults and older adults with hearing loss might be attributable to a number of possible causes.

1.3.1.2. Children and the noisy classroom

Older adults, however, are not the only group who face difficulties with multi-talker listening. Children with hearing loss face the challenge of attending to a teacher when other students are talking (Kochkin, 2005). They may be at a particular disadvantage because they are trying to learn language in classrooms that are typically noisy.

Lundeen (1991) estimates that more than 1 in 40 school children in the United States have hearing loss that interferes with their education. Children with mild-to-moderate cochlear hearing loss have delayed vocabulary compared to normally-hearing children of the same age (Pittman, Lewis, Hoover, & Stelmachowicz, 2005). Halliday and Bishop (2005) administered a battery of language and literacy tests to children with mild-to-moderate hearing loss and to age-matched controls with normal hearing. All participants were between 6 and 13 years of age. Children with hearing loss scored more poorly than controls at word reading and non-word repetition. Their scores showed significant correlations with frequency discrimination accuracy at 1 and 6 kHz. This result suggests that even mild or moderate hearing loss interferes with a child’s language development. Moreover, because some of the children were receiving speech and language therapy, the results may underestimate the effects of hearing loss on language development.

1.3.1.3. Hearing aids and multi-talker listening

Most individuals with moderate or severe hearing loss are able to receive acoustic hearing aids from the NHS. Acoustic hearing aids amplify the signal that reaches the ear. They consist of several parts. The microphone converts sounds in the environment into digital signals, the amplifier adjusts the level of these signals, and the receiver converts the amplified signal into sounds that are fed into the ear canal. Implementation of these devices has been successful and typically leads to improvements in patients' ability to understand speech and detect sounds in quiet settings (Marrone et al., 2008a).

In everyday life, however, listeners often encounter speech in the presence of background noise. Many hearing-aid users complain that listening in noise is exhausting (Edwards, 2007). Modern acoustic hearing aids implement direction-sensitivity that suppresses sounds from directions other than straight ahead. This processing may help listeners to detect sounds coming from directly in front of them in the presence of noise sources that originate from other directions. Consistent with this idea, acoustic hearing aids have been shown to provide a small benefit on self-report measures of speech intelligibility in noisy environments (Gatehouse & Akeroyd, 2006; Noble & Gatehouse, 2006; Noble, 2006). For example, Noble and Gatehouse (2006) administered the SSQ to patients who were awaiting amplification and those who had been using acoustic hearing aids for 6 months. They found that patients with hearing aids reported better intelligibility than patients without hearing aids in a variety of contexts, including speech identification in quiet, in groups, and in noise. Nevertheless, noisy environments still remain a problem for listeners who use acoustic hearing aids and is a salient factor that contributes to perceived disability (Gatehouse & Noble, 2004).

1.3.2. Encoding of sounds in the impaired ear

Sensorineural hearing loss indicates impairment of function inside the cochlea, such as outer and inner hair cells, and atypical processing at neural structures outside of the cochlea, such as the auditory nerve (Moore, 2007). These impairments have multiple consequences for the processing of speech. To fully understand the difficulties faced by hearing-impaired listeners in multi-talker environments, it is necessary to consider the peripheral processes that contribute to hearing loss. The most obvious symptoms of hearing loss are elevated thresholds for detecting weak sounds (Moore, 2007). In addition, hearing-impaired listeners show impaired frequency selectivity and also, within frequency bands, impaired representations of temporal fine structure. It is likely that all of these factors contribute to poorer segregation of the components of competing talkers and, thus, contribute to poorer speech intelligibility than listeners with normal hearing (Festen & Plomp, 1990). The following sub-sections will briefly overview these peripheral factors and their possible consequences for speech intelligibility.

1.3.2.1. Reduced sensitivity

Listeners with hearing loss have higher audiometric pure-tone thresholds than normally-hearing listeners. Loss of function of the outer and inner hair cells can contribute to elevated thresholds by affecting some of the processes involved in transducing sounds from the cochlea to the auditory nerve (Figure 1.8). For example, damage to the outer hair cells affects the active mechanism, which means that weak sounds have lower than normal amplitude on the basilar membrane. Therefore, the amplitude of a sound must be larger in order to produce detectable vibrations on the basilar membrane for hearing-impaired than for normally-hearing listeners. The inner hair cells are responsible for converting mechanical vibrations on the basilar membrane into neural activity at the auditory nerve. Damage to the inner hair cells reduces the efficiency of conversion, which means that the amplitude of basilar membrane vibrations must be larger than normal to reach threshold levels of neural activity. Moore (2007) suggests that, although the pattern of damage varies between individuals, moderate hearing loss results from damage mainly to the outer hair cells, whereas patients with severe hearing loss might also have damage to the inner hair cells.

One implication of higher detection thresholds is that hearing-impaired listeners might not hear the weaker sounds in speech, such as the sounds produced by the spoken consonants ‘p’, ‘t’, and ‘k’ (Moore, 2007). Overall, the proportion of the speech spectrum that is above threshold will be lower for hearing-impaired listeners than for normally-hearing listeners, which leads to poorer speech intelligibility (Lee & Humes, 1993; Zurek & Delhorne, 1987). When listening in background noise, normally-hearing listeners show higher accuracy of speech intelligibility when the masking sound contains temporal ‘dips’, or ‘glimpses’, where the SNR is higher than at other parts of the speech signal (Cooke, 2006). However,

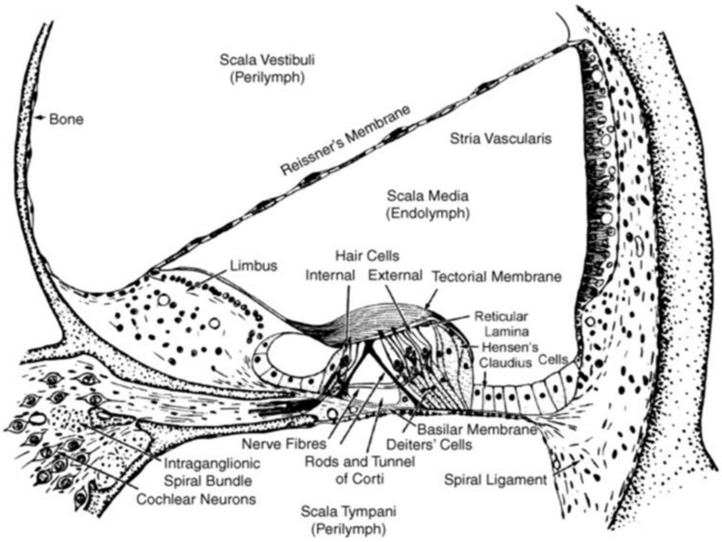


Figure 1.8. From Moore (2007). Cross-section of the cochlea, showing the basilar membrane and Reissner’s membrane.

hearing-impaired listeners show less improvement from temporal fluctuations than normally-hearing listeners (Bronkhorst & Plomp, 1992; Duquesnoy, 1983). The accuracy of speech intelligibility for hearing-impaired listeners in a temporally-fluctuating masker correlates negatively with average thresholds in quiet (Gregan, Nelson, & Oxenham, 2013). It has been suggested that higher detection thresholds may mean that the 'dips' that are detected by hearing-impaired listeners might not be able to be utilised if much of the target speech remains below threshold (Gregan et al., 2013).

There is increasing evidence that, although higher detection thresholds contribute to poorer speech intelligibility, suprathreshold discrimination abilities also affect the accuracy of speech intelligibility in background noise (Arbogast, Mason, & Kidd, 2005; Dreschler & Plomp, 1985; Tyler, Summerfield, Wood, & Fernandes, 1982). This evidence will be discussed in the following sections.

1.3.2.2. Frequency selectivity

There is a well-established relationship between thresholds for identifying speech in noise and measures of frequency selectivity (Festen & Plomp, 1983). Frequency selectivity is the ability to represent fluctuations in energy in one frequency band, separate from fluctuations in different frequency bands. In normally-hearing listeners, fluctuations in specific frequency bands are represented at specific places along the basilar membrane and, provided the frequency separation is sufficiently large, in different fibres of the auditory nerve. This type of representation is known as 'place coding'. Damage to the outer hair cells distorts the place coding system—tuning on the basilar membrane is less sharp than normal (Gaudrain, Grimault, Healy, & Béra, 2007; Moore, 2008a). One consequence of this distortion is spectral smearing of components of sounds: adjacent frequency components, which are coded into separate frequency channels in the normally-functioning auditory system, are represented in the same channel in a system with degraded frequency selectivity (Darwin, 2008).

Broad tuning predicts that greater amounts of noise will enter each auditory filter, leading to poorer SNRs at affected frequencies (Darwin, 2008; Gaudrain et al., 2007). Consistent with this idea, Marrone *et al.* (2008b) proposed that reduced frequency selectivity leads to greater energetic masking. Indeed, Arbogast *et al.* (2005) found greater energetic masking in hearing-impaired listeners than normally-hearing listeners. Participants identified speech in different types of background noise: different-band speech, different-band noise and same-band noise. The masker level was adjusted for each participant to equate sensation level. For the different-band noise masker, listeners with mild-to-moderate hearing loss showed greater SRTs than normally-hearing listeners. Listeners with hearing loss also displayed a greater shift from SRTs measured in quiet to thresholds measured in different-band noise. These results are consistent with the idea of broader auditory filters in listeners with hearing loss.

Baer and Moore (1994) processed speech by spectral smearing, aiming to simulate one aspect of hearing loss for normally-hearing participants. They found that the level of smearing affected the intelligibility of a sentence spoken in the presence of a competing talker. Intelligibility for sentences presented with three-factor broadening decreased more steeply with decreasing SNR than was found using unprocessed sentences. When six-factor broadening was used, intelligibility was low at all SNRs. These simulation results demonstrate that spectral smearing has the potential to negatively affect intelligibility in hearing-impaired listeners.

1.3.2.3. Encoding temporal fine structure

A waveform can be decomposed into two components: slow envelope modulations, and rapid modulations of temporal fine structure (TFS; Moore, 2008b). Hearing-impaired listeners have problems encoding TFS information. For example, Lorenzi, Gilbert, Carn, Garnier, and Moore (2006) processed spoken consonants using the Hilbert transform (Hilbert, 1912), with the aim of preserving *either* the speech envelope or TFS. Participants had to identify the consonant that was presented on each trial. Normally-hearing participants were able to identify the processed consonants with high accuracy using envelope cues or TFS alone. Participants with moderate hearing loss were able to identify consonants with envelope cues almost as accurately as normally-hearing participants. However, they performed very poorly at identifying consonants with only TFS information. This finding suggests a specific deficit in extracting TFS information from speech.

The underlying cause of atypical TFS encoding is unclear, although several possibilities have been suggested (see Moore, 2008b for a summary). One possibility is that broader auditory filters lead to more complex TFS information, which makes it more difficult to decode (Moore, 2008b). In more detail, broader auditory filters would lead to TFS information, at a given frequency, that was more complex and more rapidly-varying than normal. Deficits at this stage would put pressure on the mechanisms that ‘decode’ TFS information, which might not be able to track such rapid changes in frequency (Moore & Sek, 1996). Another possibility is that disruption of the waveform at the basilar membrane interferes with the cross-correlation mechanism thought to be important for decoding TFS information (e.g. Carney, Heinz, Evilsizer, Gilkey, & Colburn, 2002; Loeb, White, & Merzenich, 1983). Carney et al. (2002) express their cross-correlation model in terms of coincidence detector cells that receive inputs from the auditory nerve, such as those in the anteroventral cochlear nucleus. The model compares temporal response patterns across auditory nerve fibres that are tuned to different frequencies and assumes that the combined output of the nerve fibres provides information about spectro-temporal properties of the stimulus (Carney et al., 2002). This model suggests that precise phase-locked firing underpins accurate encoding of TFS. Moore (2008b) suggests that changes in the relative phases of a waveform at different points on the basilar membrane in hearing-impaired listeners could interfere with the cross-correlation mechanism. The cross-correlation mechanism could be distorted as a result of broader frequency tuning in individual nerve fibres and by atypical temporal responses at the basilar membrane. It is possible that disruption of the cross-correlation mechanism, broader auditory filters, or both of these processes might contribute to atypical TFS encoding in hearing-impaired listeners.

TFS is thought to be important for listening in the ‘dips’ of a masking sound (Moore & Glasberg, 1987). When only envelope cues are present, even normally-hearing participants display poor intelligibility in the presence of a fluctuating background sound (Qin & Oxenham, 2003; Zeng et al., 2005). Therefore, one reason that hearing-impaired listeners might find noisy situations difficult is that they get fewer ‘glimpses’ of the target signal as a result of reduced TFS information. Lorenzi *et al.* (2006) found that consonant identification for stimuli containing only TFS correlated with participants’ ability to take advantage of the temporal dips of background noise when identifying unprocessed speech. Furthermore, reduced audibility in the ‘dips’ (Gregan et al., 2013) may make it more difficult for hearing-impaired listeners to utilise the ‘glimpses’ that they do hear (Section 1.3.2.1; Bronkhorst & Plomp, 1992; Shinn-Cunningham & Best, 2008).

1.3.3. Ability to use cues for segregation

Listeners with normal hearing are able to use differences in spatial location and F0 between simultaneous talkers to improve intelligibility (Bronkhorst & Plomp, 1988; Brungart, 2001). However, hearing-impaired listeners typically show less improvement from separating talkers in spatial location or F0 (Arehart, King, & McLean-Mudgett, 1997; Marrone et al., 2008b). There are at least two possible explanations for this finding: (1) atypical peripheral encoding of speech, which leads to a lack of discriminability of differences in location/F0, or (2) atypical selective attention to speech based on cues for location/F0. It is likely that both of these processes interact to produce the deficits observed in noisy environments. However, I first consider the former possibility. This section draws on research discussed in the previous section to explore how atypical peripheral processing might cause difficulties discriminating location and F0 differences and, consequently, difficulties using these cues to segregate simultaneous talkers.

1.3.3.1. Ability to utilise differences in fundamental frequency

Previous studies report that hearing-impaired listeners gain smaller improvements in intelligibility than normally-hearing listeners when two simultaneous voices are separated in F0 compared to when the talkers have the same F0. For example, Arehart *et al.* (1997) presented simultaneous vowel sounds to normally-hearing and hearing-impaired listeners. F0 separation was varied from 0 to 8 semitones and participants were asked to identify one of the two vowels. SNR was manipulated in order to identify SRTs in listeners with normal hearing and listeners with moderate-to-severe sensorineural hearing loss. Hearing-impaired listeners received a mean benefit of 4.4 dB SNR across all F0 separation values compared to when the target and masker had the same F0. However, normally-hearing listeners benefited from F0 separation more than twice as much (9.4 dB). Mackersie *et al.* (2011) found similar results for spoken sentences. Listeners with moderate hearing loss showed a significant improvement in intelligibility when talkers were separated by F0s of 6 or more semitones compared to 0 semitones. However, while normally-hearing listeners received an additional benefit from increasing F0 separation to 9 semitones, hearing-impaired listeners did not (Figure 1.9A). Additionally, hearing-impaired listeners showed no benefit from separating talkers in vocal tract length (Figure 1.9B). Together, these findings suggest that hearing-impaired listeners are unable to take full advantage of talker and gender differences when listening to two simultaneous talkers. These results are consistent with the finding that hearing-impaired adults are less accurate than normally-hearing adults at discriminating between different voices, for example when asked to identify a female target talker out of three consecutively-presented female talkers (Helfer & Freyman, 2008).

There are several reasons why discriminating talker identity might be difficult for hearing-impaired listeners. First, differences in F0 might be more difficult to detect due to impaired frequency resolution. In normally-hearing listeners, the harmonics of speech spoken by talkers that are separated in F0 excite different positions on the basilar membrane. In hearing-impaired listeners, because place coding is less precise than in normally-hearing listeners, the excitation patterns associated with voices that are separated in frequency may still overlap on the basilar membrane, thus resulting in poorer accuracy for talker discrimination. Consistent with this idea, the degree of masking between two talkers that are separated in F0 increases as frequency selectivity decreases (Darwin, 2006).

Another factor that might contribute to difficulty discriminating talker identity is poorer encoding of TFS. An impaired ability to encode or interpret TFS reduces the ability to

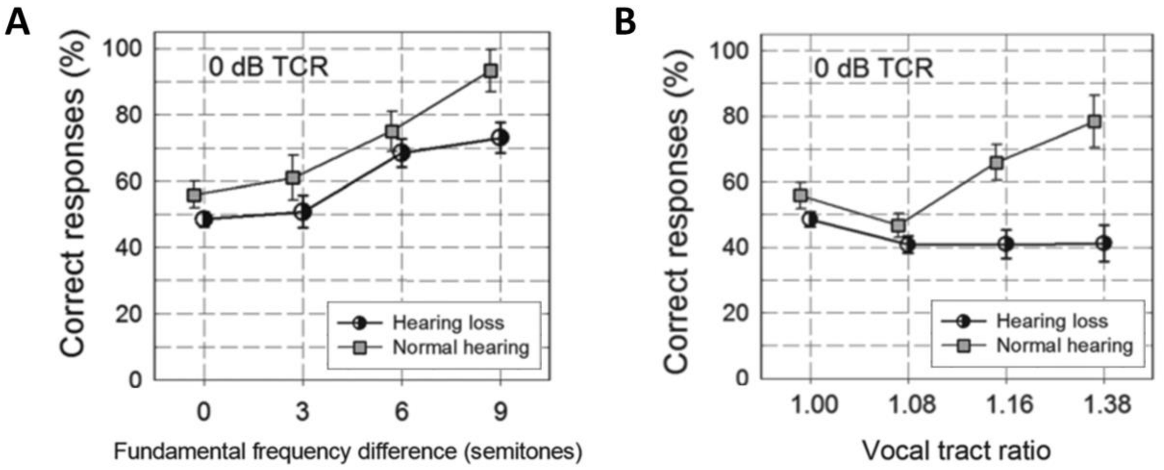


Figure 1.9. Adapted from Mackersie *et al.* (2011). Graph **A** shows the percentage of correct key word identifications for participants with moderate sensorineural hearing loss (circles) and normal hearing (squares). Percent correct is plotted as a function of fundamental frequency (F0) difference in semitones at 0 dB target-to-masker ratio (TCR) when the target was the higher-F0 talker. Graph **B** shows the percentage of correct key word identifications as a function of difference in vocal tract ratio (VTR) at 0 dB TCR, averaged over the higher- and lower-VTR talkers. Error bars on both graphs indicate ±1 standard error.

determine the pitch of complex sound (e.g. Leek & Summers, 2001; Moore, 2008b). Hence, given that differences in pitch provide important cues for segregating concurrent talkers, impaired processing of TFS will give rise to particular difficulties in multi-talker listening. Taken together, impaired frequency selectivity and encoding of TFS information might undermine the ability of hearing-impaired listeners to use F0 and pitch as cues to segregate speech spoken by two simultaneous talkers.

1.3.3.2. Ability to utilise differences in spatial location

Similar to differences in F0, increasing the spatial separation of a target talker from interfering talkers increases the accuracy of speech intelligibility for hearing-impaired listeners, but not as much as for normally-hearing listeners (Marrone et al., 2008a, 2008b; Neher et al., 2009). For example, Marrone *et al.* (2008b) presented participants with three simultaneous sentences. The target sentence was presented from straight ahead (0° azimuth) and the maskers were either collocated with the target or spatially-separated at ± 90° azimuth. Normally-hearing participants (old and young) were able to report key words from the target sentence more accurately when maskers were spatially-separated than when they were collocated. However, listeners with mild-to-moderate sensorineural hearing loss received a smaller benefit of spatial separation than normally-hearing listeners, with some listeners showing no benefit above the spatially collocated condition. The improvement in accuracy gained from spatial separation negatively correlated with thresholds for identifying speech in quiet. The results of Marrone *et al.* show that hearing-impaired listeners are not able to fully utilise spatial separation to improve intelligibility.

Marrone, Mason, and Kidd (2008c) investigated the conditions in which normally-hearing listeners benefit from spatial separation in three-talker listening. They found significant spatial release from masking when the maskers were located at ± 15° from the target. This result is consistent with previous research that shows spatial release from masking in normally-hearing listeners (e.g. Allen et al., 2008; Noble & Perrett, 2002). However, when listeners wore an earplug and earmuff over one ear, spatial release from masking was nearly eliminated. This finding suggests that unilateral hearing loss might distort, or eliminate, the binaural cues necessary to benefit from spatial separation.

The aforementioned results are consistent with the finding of impaired spatial localisation abilities in hearing-impaired listeners (Noble, Byrne, & Ter-Horst, 1997). Reduced ability to extract TFS information in hearing-impaired listeners can lead to impaired ability to encode ITDs, which are important cues for sound localisation (Hawkins & Wightman, 1980; Lacher-Fougère & Demany, 2005). Bronkhorst and Plomp (1988) investigated the relative contributions of ITDs and ILDs to speech intelligibility in the presence of speech-shaped noise. They manipulated the noise to contain only ITD or ILD information and simulated spatial azimuths of 0°, 30°, and 90°. Speech was presented from the front (0°) and SNR was varied in order to measure SRTs. Listeners with symmetrical mild-to-moderate sensorineural hearing loss showed a similar improvement in thresholds as normally-hearing listeners when ITDs were present compared to when ITDs were not present. The improvement in thresholds when ILDs were present varied among hearing-impaired listeners, ranging from 0 dB SNR to greater than 7 dB SNR. This variation was partially explained by pure-tone thresholds at 4 kHz at the ear contralateral to the noise source. Bronkhorst and Plomp propose that this finding is consistent with the explanation that the head shadow effect at 90° is most prominent at 3–5 kHz. This result suggests that high-frequency hearing loss might reduce the ability to benefit from ILDs for separating speech from broadband noise. It is likely that this factor also undermines the ability of hearing-impaired listeners to segregate spatially-separated talkers during multi-talker listening.

Consistent with the idea that the results of Marrone *et al.* (2008b, 2008c) do not result from reduced audibility, acoustic hearing aids do not greatly improve the ability to benefit from spatial separation. Marrone *et al.* (2008a) reported a significant but small speech intelligibility benefit of bilateral hearing aids over no hearing aids when three talkers were separated (-90°, 0°, and +90°) compared to collocated (0°). Even with hearing aids, hearing-impaired listeners showed substantially less improvement in speech intelligibility from spatial separation than normally-hearing listeners. This finding is consistent with the fact that acoustic hearing aids amplify sounds, rather than correcting for reduced ability to extract or encode TFS information. It has been suggested that, in some cases, acoustic hearing aids may even decrease intelligibility compared to unaided listening because the head shadow is reduced at the microphone of a hearing aid (Bronkhorst, 2000).

1.3.4. Summary and conclusions

This section has highlighted the difficulties posed by multi-talker environments for listeners with sensorineural hearing loss. These difficulties can lead to social isolation and problems acquiring language in childhood. Even when using an acoustic hearing aid that improves speech recognition in quiet, there is little or no benefit of a hearing aid for intelligibility during multi-talker listening. This outcome highlights that, to be successful, rehabilitation not only needs to ensure that talkers are audible, but also that patients are able to segregate simultaneous talkers and attend selectively to a talker of interest.

Unlike normally-hearing listeners, hearing-impaired listeners do not receive as much speech intelligibility benefit from separating simultaneous talkers in spatial location or F0. This finding might partially result from the distorted encoding of sounds at the periphery. Several peripheral factors might contribute, including reduced sensitivity, broadening of auditory filter bandwidths, and reduced encoding of TFS.

It is possible that central, in addition to peripheral, factors contribute to poorer speech intelligibility by hearing-impaired listeners. However, the study of older adults with hearing loss in many of the experiments discussed in this section limits our estimation of the extent to which the cognitive consequences of hearing loss contribute to difficulties in multi-talker listening relative to general cognitive decline in older adults. In addition to hearing loss, older adults experience cognitive decline that is independent of hearing loss itself (although may be correlated with it). This decline might involve factors such as reduced working memory capacity (Braver & West, 2008) and reduced speed of processing (Salthouse, 1996). Thus, it is difficult to distinguish whether difficulties with multi-talker listening in older adults result from the peripheral and cognitive factors associated specifically with hearing loss or an aspect of cognitive decline that is not a direct result of hearing loss. Studying children with hearing loss could help to isolate cognitive factors that are associated specifically with hearing loss, independent from the effects of older age on cognition.

The next section will consider the evidence that is available for one cognitive factor that might be associated with hearing loss: reduced ability to *attend* selectively to a talker of interest. If hearing-impaired listeners show atypical cognitive processing relative to normally-hearing listeners, then understanding the central consequences of hearing loss is likely to be important for rehabilitation.

1.4. Auditory attention in hearing loss

In recent years, there has been an explosion of interest in the relative contributions of peripheral and central factors to the difficulties faced by people with hearing loss, within both clinical and scientific communities. This idea recognises that problems at the ear might have upstream consequences for central processing. Central processing may be particularly important for speech intelligibility when listening in background noise, where there are multiple concurrent sounds that need to be segregated. It has been suggested that difficulties with attention might explain why listeners who use acoustic hearing aids find communication in quiet relatively easy compared to multi-talker listening (Ihlefeld, Shinn-Cunningham, & Carlyon, 2012).

There are several reasons to believe that the attentional processing of speech might be atypical in listeners with hearing loss. This section will start by discussing this rationale. Although several researchers have speculated about atypical attentional processing, there has been little experimental evidence to support their speculations. Peripheral and central factors that contribute to poor intelligibility are notoriously difficult to tease apart. This section will discuss some experiments that have attempted to explore central factors. I will argue that previous experiments have not been rigorous in ruling out the explanation that differences in peripheral, rather than central, processing may explain differences between normally-hearing and hearing-impaired listeners. At the end of this section, some possible directions for future research will be suggested.

1.4.1. Rationale for atypical auditory attention

It is well-accepted amongst clinicians and scientists that there is wide variability in the accuracy of speech intelligibility that is unrelated to audiometric thresholds (Grose & Hall, 1996; Neher et al., 2009) and frequency selectivity (Rose & Moore, 1997). Neher *et al.* (2009) also found that the ability of hearing-impaired listeners to benefit from spatial separation varied among listeners with similar audiometric thresholds and identical acoustic hearing aid fittings. The audiogram only provides information about the early stages of auditory input (Kraus & Anderson, 2012). Therefore, it is possible that individual differences may result from suprathreshold differences in peripheral encoding or from differences in central processing.

Shinn-Cunningham and Best (2008) assumed that atypical peripheral processing contributes to failures in later stages of processing, such as the ability to attend selectively to acoustical stimuli. This idea followed from the common conceptualisation of the auditory system as a sequence of processing stages that build upon each other (Moore, 2007). Shinn-Cunningham (2007) assumed that there are two possible points at which selective attention can fail. The first is the ability to segregate the target from maskers, suggesting that a failure in object formation or, in other words the grouping of spectro-temporal features, is the cause of difficulty. Failures at this stage might arise when the spectro-temporal features of the target are not easily distinguishable from features of the maskers. Such failures are more likely to occur in hearing-impaired listeners than normally-hearing listeners due to reduced spectro-temporal resolution. If spectro-temporal features are not segregated effectively, then interference between a target and masker is more likely, thereby reducing intelligibility.

The second point at which selective attention might fail is in selecting the correct object to which to direct attention. This idea implies that the target might be effectively separated from the maskers, but the listener incorrectly attends to a masker rather than the target. Failures in object selection might also arise from spectro-temporal similarity between the target and maskers, in this case causing confusion about which object is the target. Alternatively, the listener may not know to which feature to direct attention, which might arise when higher-order cues for segregation, such as location or pitch, are not salient or when the listener does not have experience of selecting a particular feature. Another potential difficulty is that the maskers might be inherently more salient than the target and the listener cannot overcome distraction from the masker(s). The latter might result from increased masking in hearing-impaired listeners due to broader auditory filter bandwidths.

A visual analogy is an observer who is colour blind (Figure 1.10A). If the observer does not have the ability to detect differences between red and green objects, then asking the observer to pay attention to the red object may be problematic for two reasons. First, a lack of discriminability might impair their ability to segregate the red and green objects if the objects are also similar on other dimensions, such as the case where they overlap in spatial location (Figure 1.10B). Therefore, they will not be able to pay attention to the red object because they do not perceive it as an object. Second, even if the observer can discriminate the red and green objects enough to segregate them, the observer might not be able to select one object for attention because colour is not a salient cue and they may not be certain which object is red (Figure 1.10C).

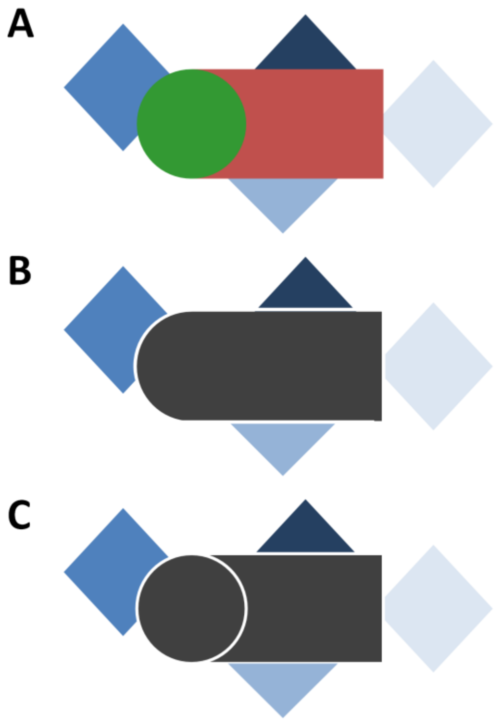


Figure 1.10. Diagrams to visualise the analogy of an observer who is colour blind. (**A**) A simplified example of visual objects, exemplified by shapes of different colours that overlap in spatial location. (**B**) An example of difficulties with object formation. The red and green objects from panel **A** are perceived as a single object. (**C**) An example of difficulties with object selection. The red and green objects from panel **A** are segregated, however, based on colour information alone, it is unclear which of the two objects should be selected for attention.

These two hypotheses make different predictions for hearing-impaired listeners' performance on multi-talker tasks. If object formation is the problem, then incorrect responses are likely to consist of a mixture of words spoken by the target and/or maskers. If object formation is intact but object selection is the problem, then incorrect responses are more likely to consist of words spoken by one of the maskers. It is possible that both of these mechanisms are impaired in listeners with hearing loss, although the factor that is most problematic might depend on the structure of the acoustical environment. Although researchers have speculated that both of these processes might be impaired (e.g. Best et al., 2009; Neher et al., 2009; Shinn-Cunningham, 2008), no previous experiments have directly investigated these hypotheses.

1.4.2. Cognitive factors correlate with performance

One line of research has investigated the relationship between cognitive factors and thresholds for identifying speech in noise through correlation analyses. For example, Neher *et al.* (2009) measured SRTs in hearing-impaired listeners when two speech maskers were spatially-separated to the left and right. Thresholds significantly correlated with selective attention (Test of Everyday Attention [TEA] sub-test 1), attentional switching (TEA sub-test 4), and working memory (reading span test). Parbery-Clark, Skoe, Lam, and Kraus (2009) found that musicians, who show better-than-average SRTs in noise, also have above-average working memory capacity (Parbery-Clark et al., 2009). Furthermore, SRTs in musicians correlated with working memory capacity. In addition to SRTs, working memory capacity covaries with the perceived effort of listening in noise in normally-hearing adults (Rudner, Lunner, Behrens, Thorén, & Rönnberg, 2012).

Together, these results provide evidence for a link between cognitive factors, such as attention and working memory, and SRTs and effort of listening in noise. However, the direction of this relationship is not clear. A causal relationship cannot be established between cognitive ability and SRTs based on these experiments. Furthermore, most of Neher *et al.*'s (2009) participants were older adults, which means that it is difficult to distinguish between changes in cognitive factors resulting from hearing loss and changes resulting from older age.

1.4.3. Cognitive training

If cognitive factors, such as attention, contribute to difficulty identifying speech in noise, then training participants on aspects of cognition relevant to processing speech might improve speech intelligibility in noise. Song, Skoe, Banai, and Kraus (2012) trained a group of normally-hearing participants who displayed large individual differences in accuracy of speech intelligibility in noise. Training included a variety of tasks from the 'Listening and Communication Enhancement' program (Neurotone, Inc., 2005), including practice with speech-in-noise, speech masked by a competing talker, time-compressed speech, auditory closure, and auditory memory. Training was provided for twenty 30-minute sessions across four weeks. The group that had undergone training showed better SRTs in noise than the control group who had received no training. Furthermore, the improvement in thresholds was retained six months later, which demonstrates that the effects of training were sustained rather than transient. The results of Song *et al.* show that training has the potential to improve speech intelligibility in background noise.

However, there are two features of many previous training studies that limit the conclusions that can be drawn. First, Song *et al.* (2012) compared training with a control group who did not undergo any training. Therefore, it could be argued that other differences between the groups, such as differences in motivation or familiarity with the testing equipment, might explain differences in performance. Second, the specific tasks that improve SRTs remain unclear, since Song *et al.* employed a variety of tasks. Even within a single task, it is not clear which aspect of the task caused improvements in intelligibility. Improvement might be the result of any aspect of familiarity with the task, which includes peripheral acoustical processing as well as central auditory processing (Amitay, Zhang, Jones, & Moore, 2014).

Furthermore, not all studies have found a significant effect of training on speech intelligibility. For example, Burk and Humes (2007) provided training to young, normally-hearing listeners. During training, words that were described as lexically-difficult (defined as high neighbourhood density and low frequency of usage) were spoken by a single talker in the presence of speech-shaped noise. Participants had to identify the presented word from an alphabetical list of 75 possible words. Training was presented for between 8 and 11 hours, spread over two-week period, and participants received orthographic feedback regarding their responses. The test phase presented the same lexically-difficult words and a different set of lexically-easy words. In the open-set condition, participants had to identify the spoken words. The closed-set condition was identical to the task during training. The accuracy of reporting the familiar lexically-difficult words was higher after training than before training. This improvement generalised to new lexically-easy words in the closed-response condition, but not to lexically-easy words in the open-response condition. Similar patterns of results were found when training was extended to up to 25 hours. These results show limited benefits of training that are restricted to the words and task presented during training. Therefore, this type of training might be of limited use for improving the accuracy of identifying speech in noise during everyday listening.

Overall, the possible benefits of training for speech intelligibility in noise are unclear. This uncertainty arises from mixed findings, the implementation of no-training comparison groups, and the difficulty of elucidating the causes of speech intelligibility benefits following training. Nevertheless, the results from Song *et al.* suggest that training has the potential to improve speech intelligibility in noise. One possible explanation of the results of Song *et al.* is that training improved cognitive processes, such as attention.

1.4.4. Preparatory attention in hearing loss

During multi-talker listening, there is some evidence that hearing-impaired listeners (unlike normally-hearing listeners) receive only a small benefit from advance cues that provide information about the target talker, such as their spatial location. For example, Gatehouse and Akeroyd (2008) asked older adults with hearing loss to identify words spoken by one talker in the presence of two competing talkers and speech-shaped noise. Light-emitting diodes (LEDs) attached to the loudspeakers cued the spatial location of the target, the time the target would start talking, or both attributes. The results showed a significant but small (2%) improvement in the accuracy of speech identification in the cued than the no-cue condition. However, there was large individual variability. Only 16 out of the 57 participants (28%) showed an average improvement of 5% or more in cued conditions than the no-cue condition. This finding demonstrates that some, but not all, hearing impaired listeners can use advance cues to slightly improve the accuracy of speech intelligibility.

Best *et al.* (2009) showed that the intelligibility benefit of receiving advance cues was smaller in hearing-impaired listeners than in normally-hearing listeners. Participants were asked to report a digit sequence spoken by a target talker. On each trial, the target speech occurred in one of five time windows and at one of five spatial locations. Maskers, which consisted of time-reversed speech, were presented during other time windows at the target's spatial location and at the other spatial locations during the target time window. There were four different cueing conditions: no cue, 'when' cue (indicating the time window during which the target speech would occur), 'where' cue (indicating the spatial location from which the target speech would occur), and 'both' cue (indicating the time window *and* spatial location of the target speech). The sensation level of the target speech was equated for normally-hearing and hearing-impaired listeners, which produced similar mean accuracy in the no-cue condition across both groups of listeners. The intelligibility benefit of knowing the spatial location of the upcoming talker (calculated as the difference in percent correct between the 'where' cue condition and the no-cue condition) was significantly smaller for hearing-impaired listeners than for normally-hearing listeners. However, hearing-impaired and normally-hearing listeners received similar gains in intelligibility from the 'when' cue compared to the no-cue condition. This result is consistent with the hypothesis that hearing-impaired listeners do not utilise preparatory *spatial* attention to improve the accuracy of speech intelligibility to the same extent as normally-hearing listeners. However, a possible alternative explanation, which is not addressed by Best *et al.*, is that hearing-impaired listeners require more time than normally-hearing listeners to effectively prepare their attention for the location of an upcoming talker.

Little research has explored the ability of hearing-impaired listeners to use prior knowledge of a talker's gender or F0 to improve speech intelligibility. One related experiment explored the effect of voice familiarity on intelligibility in older (> 60 years old) listeners (Johnsrude et al., 2013). Participants heard two simultaneous sentences on each trial and had to report key words spoken by a target talker. Intelligibility was significantly higher when either the target or distracting voice was the participants' spouse than when both voices were unfamiliar. This finding suggests that familiarity with a voice improves speech intelligibility for older listeners.

One question that remains is whether *young* hearing-impaired listeners also receive an intelligibility benefit of knowledge of a talker's gender or F0 because Johnsrude *et al.*'s (2013) results may reflect a greater reliance on voice experience with older age. Investigating the effects of prior knowledge is an interesting direction for future research because differences in cognitive ability, independent from peripheral differences, can be explored.

1.4.5. Summary and conclusions

Overall, there are several reasons why hearing-impaired listeners might show atypical attention to speech. Impaired peripheral processing may reduce the ability to segregate simultaneous talkers and to receive an improvement in the accuracy of speech intelligibility from separating simultaneous talkers in spatial location or F0. However, sensorineural hearing loss might also have central consequences, including the ability to direct attention to target speech and ignore distracting noise.

Two branches of research imply a link between cognitive processing and SRTs in noise. First, SRTs correlate with cognitive factors, such as attention and memory. Second, there is some evidence that SRTs in noise can be improved with training that provides practice identifying speech in noisy environments. However, the conclusions that can be drawn from either of these approaches are limited by the lack of causal inference.

Cueing the location of an upcoming target talker does not improve the accuracy of speech intelligibility for hearing-impaired listeners as much as for normally-hearing listeners. Little research has investigated this idea and several questions remain, such as whether cueing gender or F0 improves speech intelligibility for hearing-impaired listeners.

1.5. Overall conclusions

Speech is typically encountered in the presence of background noise, including the voices of other competing talkers. In these situations, normally-hearing listeners benefit from improved intelligibility when competing talkers are separated from the target in spatial location or F0. Furthermore, intelligibility is improved if listeners are given time to prepare for the location or F0 of an upcoming talker than when they do not know these attributes in advance. The time course of preparatory attention for multi-talker listening is unclear. There are two possibilities: (1) the time interval does not influence intelligibility until it reaches a certain threshold for successful preparation, beyond which increasing the time interval does not improve intelligibility further; or (2) increasing the amount of preparation time improves intelligibility progressively, at least below a certain 'optimal' interval. Understanding the time-course of preparatory attention is important for increasing knowledge of the factors that can improve speech intelligibility in challenging listening environments.

In normally-hearing listeners, directing endogenous selective attention to an acoustical stimulus activates a fronto-parietal network that either overlaps with, or is the same as, an analogous network involved in visual attention. This network controls the allocation of attentional resources to stimuli that compete for representation in sensory regions. The pattern of neural activity during multi-talker listening can be used to predict the attentional orientation of a listener with high (> 90%) accuracy.

When participants prepare for an upcoming talker, activity is observed in a similar fronto-parietal network as is observed during selective attention. The balance of activity within this network depends on whether participants know information about the spatial location or F0 of an upcoming talker. Previous experiments have revealed this network with high spatial resolution, but the time course of activity within these regions has not been mapped in detail. Knowing the time course of brain activity during preparatory attention might help improve understanding of the mechanisms that underlie the improvement in speech intelligibility when participants have time to prepare for an upcoming talker.

Listeners with hearing loss particularly struggle in noisy environments, even when they receive an acoustic hearing aid that improves their recognition of speech in quiet. One possible explanation for this finding is atypical central processing of speech, although the central consequences of hearing loss are not fully understood. Atypical peripheral processing may contribute to failures in later stages of processing, such as the ability to attend selectively to acoustical stimuli. Deficits may include the ability to separate a target from masking noise (i.e. difficulties with 'object formation') and to decide which talker is the target (i.e. difficulties with 'object selection'). These hypotheses are consistent with the finding that hearing impaired listeners report multi-talker listening to be difficult and effortful.

Experiments that have attempted to link auditory attention (by measuring performance on cognitive tests) to SRTs in noise are unable to estimate the relative contributions of peripheral and central processing to poorer SRTs. One explanation for the link between cognitive factors and SRTs in noise is that atypical central processing could, by itself, make intelligibility during speech-in-noise more challenging; for example, due to impairments in attending selectively to speech or ignoring distracting noise. Alternatively, impairments in suprathreshold peripheral processing (which are not manifest by higher audiometric pure-tone thresholds) may result in atypical central processing as a normal compensatory response to a distorted input from the periphery; for example, due to a greater perceptual load. The previous studies investigating the link between auditory attention and SRTs in noise, therefore, are unable to reveal the consequences of sensorineural hearing loss on auditory attention during speech intelligibility in noise.

The finding that young, hearing-impaired adults receive little improvement in intelligibility from knowing the spatial location of the target talker before he or she begins to speak is consistent with the idea that hearing-impaired listeners show atypical attentional processing as a result of sensorineural hearing loss. The central consequences of hearing loss, including how hearing-impaired listeners direct attention to speech in noisy environments, are not fully understood. It is important to elucidate a technique for measuring potential difficulties with auditory attention without the effects associated with atypical attention being attributable to atypical peripheral transduction. Investigating the effects of prior knowledge is an interesting direction for future research because differences in cognitive processing can be explored without the confounding effects of differences in peripheral transduction.

There are several possible explanations for the reduced benefit of prior knowledge for speech intelligibility in hearing-impaired listeners. One explanation is that hearing-impaired listeners need more time to prepare for an upcoming talker to produce an equivalent improvement in speech intelligibility as normally-hearing listeners. Another possibility is that hearing-impaired listeners do not utilise the same preparatory brain activity as normally-hearing listeners, leading to an impaired ability to utilise advance cues to improve intelligibility. Previous experiments have not investigated these hypotheses.

The experiments in this thesis first investigated the time course of brain activity in normally-hearing adults when they prepare for an upcoming talker in a two-talker (Chapter 3) and three-talker (Chapter 5) listening task. The experiments reported in Chapter 4 explored how the duration of preparation time influences the accuracy and latency of speech intelligibility in normally-hearing listeners. One aim was to devise a technique for measuring preparatory attention in normally-hearing listeners that could be applied to listeners with hearing loss. The experiments reported in Chapter 6 investigated whether the duration of preparation time affects intelligibility in hearing-impaired children and whether hearing-impaired children display preparatory brain activity. The overall aim of Chapter 6 was to investigate whether hearing-impaired listeners show atypical attentional processing of speech during multi-talker listening, while avoiding confounds due to differences in peripheral processing.

Chapter 2

Measuring the Time-course of Brain Activity with EEG

In 1929, Hans Berger showed that electrical activity could be recorded from the human brain by placing an electrode on the scalp and measuring changes in voltage over time. This technique is commonly known as electro-encephalography (EEG). An explosion of interest in cognitive neuroscience, along with the development of powerful and inexpensive computers, led to increased popularity of EEG research in the 1980s. Robust, reproducible patterns of EEG characterise many aspects of perception and cognition, including, but not limited to, attention, auditory processing, and brain pathology (e.g. Ahmadlou & Adeli, 2011; Näätänen, 1990; Niedermeyer & Lopes da Silva, 1999; Woods, Alain, Covarrubias, & Zaidel, 1993).

The main advantage of EEG compared to other neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), is high temporal resolution. Voltage changes can be measured from the scalp at the sub-millisecond scale. Many researchers now view EEG as complementary to hemodynamic measures, which have better spatial resolution but worse temporal resolution than EEG. Another advantage of EEG, along with its magnetic counterpart magneto-encephalography (MEG), is the ability to measure neural activity directly and non-invasively, rather than through indirect blood-oxygen level-dependent (BOLD) approaches.

Over the past few decades, EEG has become increasingly informative, in part due to the possibility of using larger and denser arrays of electrodes. Dense electrode arrays increase sensitivity to differences in voltage across the scalp and increase the accuracy with which inferences can be drawn about the brain regions that underlie scalp-recorded activity (Ebersole, 1997; Michel et al., 2004). For example, Lantz *et al.*, (2003) investigated the accuracy and precision with which a simulated focal source was identified using source reconstruction. By conducting simulations, they found that accuracy and precision improved as the number of electrodes increased from 25 to 100. However, there was no further improvement past 100 electrodes. They observed similar results when using EEG to reconstruct the sources of epileptogenic lesions in 14 patients. Lantz *et al*. recorded from 123 electrodes and down-sampled to 31 or 63 electrodes. The spatial correspondence between the site of epileptogenic lesion and the source reconstruction estimate improved in 9 of the 14 patients when the number of electrodes was increased from 31 to 63 electrodes. However, there was only minimal improvement between 63 and 123 electrodes. These results show that larger electrode arrays can improve the precision of source reconstruction, up to approximately 100 electrodes, after which the improvement plateaus with increasing numbers of electrodes.

In this chapter, I first discuss current understanding of the neural basis of EEG signals. The second section considers issues arising from EEG recording. Next, I discuss methods of analysing the signals that are recorded using EEG, including common processing techniques and methods for drawing statistical comparisons. This chapter concludes with a brief summary of the reasons why EEG is well-suited to addressing the aims of this thesis.

2.1. Neural basis of EEG signals

2.1.1. Activity in individual neurons

Fluctuations in voltage recorded from the scalp are thought to result primarily from the postsynaptic activity of neurons. Postsynaptic activity refers to changes in voltage across

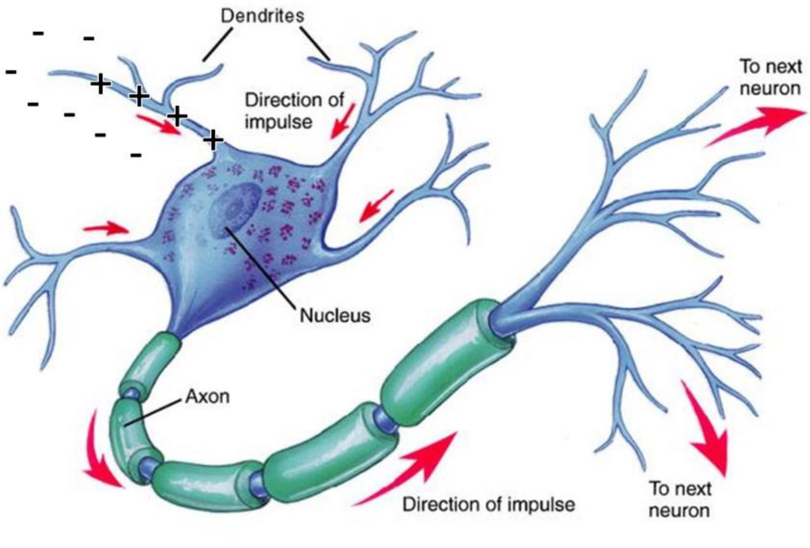


Figure 2.1. Diagram showing architecture of a neuron. Adapted from http://www.mhhe.com/socscience/intro/ibank/set1.htm [retrieved 14/08/2014]. Examples of net charges are indicated by + and - symbols.

a neuron's membrane, which results from the transmission of particles through ion channels that are situated on the membrane (e.g. opening or closing of channels due to neurotransmitter binding). These changes generate electric and magnetic fields (Nunez & Srinivasan, 2006).

In more detail, if an excitatory neurotransmitter is released from a pyramidal cell, this causes current to flow from extracellular space (i.e. outside of the neuron) into the neuron, which results in an overall negative polarity outside the neuron. To complete the circuit, current flows out of the cell body and basal dendrites of the neuron. Together, the positivity inside the neuron and negativity in extracellular space create a small dipole (i.e. a pair of positive and negative charges that are separated by distance; Figure 2.1)

2.1.2. Synchronous activity

The electrical activity of a single neuron cannot be measured at the scalp. One reason is that the electrical distance[2] between neurons and the scalp is too large to detect the small currents that can be generated by a single neuron (Makeig et al., 2004). Instead, scalp potentials are believed to arise from the co-ordinated activity of populations of neurons (Luck, 2005; Rusiniak et al., 2013).

If two neurons are oriented in parallel and fire action potentials at the same time, their activity may summate, such that the voltage resulting from co-ordinated firing is twice as large as the voltage produced by a single neuron alone (Luck, 2005). If two neurons are not oriented in parallel or fire at slightly different times, then the activity of the two neurons might at least partially cancel each other out, producing a smaller combined signal. Complete cancellation will occur if the neurons are oriented at 180° and fire at the same time. Cancellation might also occur if one neuron receives excitatory neurotransmitter and an adjacent neuron receives inhibitory neurotransmitter.

It is thought that co-ordinated activity must occur at scales of several centimetres for the resulting voltage to be detectable at the scalp (Cooper, Winter, Crow, & Walter, 1965). Therefore, the voltages that are recorded on the scalp must arise from activity in large numbers of neurons that are located close together with similar orientations and co-ordinated firing (Luck, 2005; Makeig et al., 2004). Variations in the location, orientation, number, or timing of active neurons all contribute to the detailed pattern of activity that is observed on the scalp (Alain & Tremblay, 2007).

---

[2] Electrical distance is a measure of the time taken for an electromagnetic wave to travel between two locations, where greater electrical distances indicate longer durations of time.

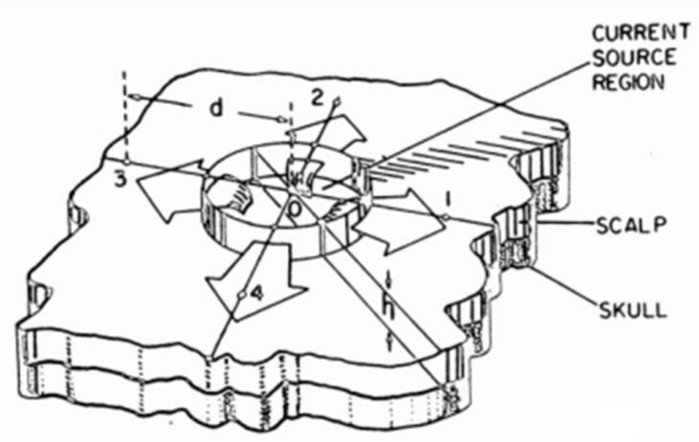


Figure 2.2. Schematic of lateral spread of activity upon reaching the scalp. From http://psdlw.users.sourceforge.net/career/dweber_docs/eeg_scd.html [retrieved 14/08/2014].

The relative orientations of neurons depend on the curvature of the cortex, which includes a number of cortical folds (Luck, 2005; Nunez & Srinivasan, 2006). The net dipole that is produced by an area of cortex is equivalent to the average of the dipoles from constituent neurons (Luck, 2005).

2.1.3. Volume conduction

Another factor that contributes to the pattern of scalp-recorded activity is volume conduction. Volume conduction results from the fact that the brain is a conductive medium, which conducts electrical activity to its surface. Current is confined inside the head, although the electrical fields extend into air surrounding the head (Nunez & Srinivasan, 2006). Since the brain is an inhomogeneous conductor, volume conduction can 'blur' the underlying activity before it reaches the scalp (Neuper & Klimesch, 2006). In particular, the scalp, skull, and cerebrospinal fluid each have different conductive properties to the brain itself (Nunez & Srinivasan, 2006). Electricity follows the path of least resistance, which means that electrical potentials are likely to spread laterally when they reach the highly-resistant skull (Luck, 2005; Figure 2.2). To complete the circuit, return currents also flow through the surrounding medium. In combination, these factors mean that activity generated in one brain region can lead to voltages on areas of the scalp that are relatively distant from the underlying source of activity (Luck, 2005).

From the opposite perspective, scalp-recorded activity at any electrode typically reflects a combination of activity generated by different neural sources whose activity overlapped in time. Although EEG measures voltage at discrete scalp locations, these locations only produce, at best, a loose indication of the location of the neural generators of

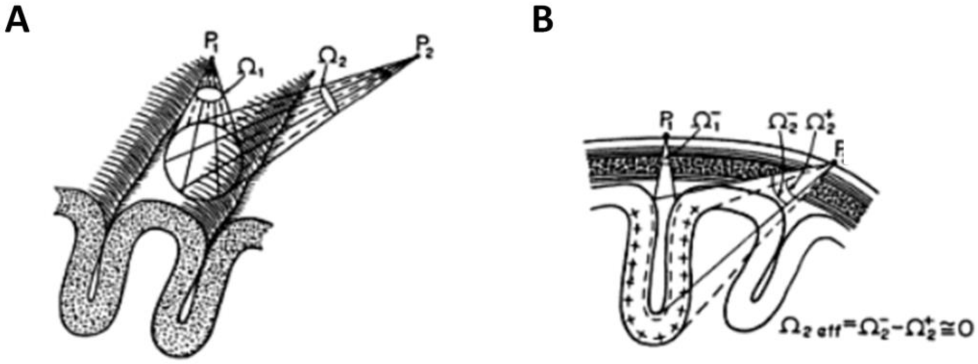


Figure 2.3. Schematic portraying the effect of source orientation on scalp potentials. (**A**) Hypothetical potentials P_1 and P_2 depend on the angles Ω_1 and Ω_2 from a hypothetical dipole later at the crown of a gyrus. (**B**) Illustration of the cancellation problem when a dipole layer occupies both walls of a sulcus. Adapted from Mountcastle (1998).

activity (Neuper & Klimesch, 2006). Estimating the location of the sources that generate scalp-recorded EEG activity is a complex problem, which will be discussed later in this chapter.

2.1.4. Selectivity of recorded activity

Due to the orientation of neurons in the brain and the placement of EEG electrodes, EEG recordings are more sensitive to certain sources than others. The electrical distance2 between neurons and the scalp is one factor that determines sensitivity. Areas that are (electrically and, often, geometrically) closer to the surface of the brain will generally make larger contributions to scalp-recorded voltages (Nunez & Srinivasan, 2006).

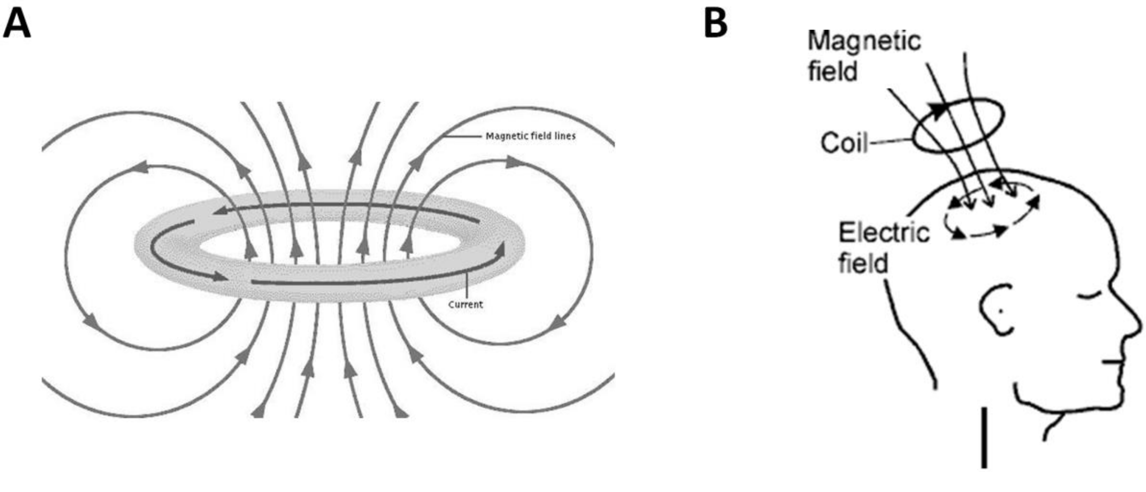


Figure 2.4. Relative orientations of electric and magnetic fields in a current coil (**A**; from http://www.askamathematician.com/2011/02/q-what-are-the-equations-of-electromagnetism-what-all-do-they-describe-to-us [retrieved 14/08/2014]) and visualised on a human scalp (**B**; adapted from http://www.biomag.hus.fi/tms/Thesis/dt.html [retrieved 14/08/2014]).

Orientation of generators is also an important factor that influences sensitivity (Figure 2.3). Pyramidal neurons (Luck, 2005) and, more generally, dipole layers that are located in the crowns of cortical gyri (Nunez & Srinivasan, 2006) are thought to make large contributions to EEG recordings because they are aligned perpendicular to the surface of the cortex. Due to the configuration of electrical fields, electrodes are most sensitive to sources aligned perpendicular and less sensitive to sources aligned parallel to the scalp. It is estimated that approximately 85% of cortical neurons are pyramidal cells oriented perpendicular to the cortical surface (Braitenberg & Schüz, 1991). Therefore, the selectivity of EEG recordings makes it well-suited to measuring potentials from groups of cortical neurons.

In addition to neuronal activity, the activity of glia are thought to contribute to EEG activity (Buzsáki, Anastassiou, & Koch, 2012). Glia have excitable membranes, which can be depolarised as a result of nearby neuronal activity. The brain contains many more glial cells than neurons and, therefore, synchronised membrane potential changes in glia have the potential to contribute to scalp-recorded electrical activity (Buzsáki, Traub, & Pedley, 2003).

The voltage field and magnetic field are oriented in perpendicular directions (Figure 2.4). Thus, EEG and MEG provide complementary information (Luck, 2005) because they are preferentially sensitive to activity in different sources. In contrast to EEG, MEG is preferentially sensitive to dipoles that are oriented parallel to the surface of the skull, such as those that sit on sulcal walls (Luck, 2005).

2.1.5. Variability of the EEG signal

Across different recording sessions, there is some variability in the EEG signal, even from the same participant. However, within-subject variability is relatively small compared to between-subject variability. One factor underlying between-subject variability is different cortical folding patterns in different individuals. In addition, there is some between-subject variability in the correspondence between specific cortical locations and functional areas. Together, these factors have the potential to give rise to differences in the location or orientation of neurons from equivalent sources. In turn, these differences affect the distribution of projected activity on the scalp (Luck, 2005; Neuper & Klimesch, 2006).

2.2. EEG recording

2.2.1. Electrical circuits in EEG recordings

Voltage is a measure of potential difference between two locations. EEG measures the voltage for current to pass between two electrodes (Luck, 2005). All EEG recordings are bipolar—that is, they measure the potential difference between two electrodes, rather than any measure at a single electrode (Nunez & Srinivasan, 2006). Although EEG signals are commonly described as taking place at a single scalp location, all EEG recordings actually measure the voltage difference between the recording electrode and a reference electrode.

2.2.2. Signal amplification

Voltage fluctuations at the scalp are typically very small, on the order of millionths of a volt (i.e. microvolts). Equipment for recording EEG amplifies the signal so that it can be measured accurately. This is important because the electrical noise that can contaminate EEG recordings is often large compared to the small voltage fluctuations that arise from neural activity.

EEG uses a differential amplifier (Luck, 2005), which involves three electrodes: a recording electrode, a reference electrode, and a ground electrode. Differential amplifiers amplify the difference between the recording-to-ground voltage and the reference-to-ground voltage. Any electrical activity recorded at the ground site is cancelled out by this amplification method, while activity at the reference site contributes to the recorded signal. The issue of choosing an appropriate reference site will be discussed later in this chapter.

2.2.3. Sampling rate

EEG amplifiers incorporate analogue-to-digital conversion, which digitises the recorded signal at a sequence of discrete time points determined by the specified sampling rate (defined as the number of samples per second).

When selecting a sampling rate for an EEG experiment, the Nyquist theorem must be taken into account. The Nyquist theorem states that the information in an analogue signal is only preserved in digitisation if the sampling rate is equal to or greater than two times the highest frequency in the signal (Luck, 2005; Nunez & Srinivasan, 2006). If the Nyquist limit is violated, then aliasing may occur, meaning that high-frequency signals might be misrepresented as low-frequency signals. In practice, Luck (2005) suggests that a sampling rate of three times the highest frequency should be used because most filters employ gradual rather than sharp cut-off rates.

2.2.4. Impedance

Impedance is the effective resistance of a circuit when voltage varies as a function of time (i.e. for alternating current sources). In EEG recordings, it is important to ensure that the impedance between the scalp and the recording electrodes is low because electricity typically follows the path of least resistance. This process usually involves abrading the outer layer of dead skin cells on the scalp and ensuring a good electrical connection by applying a conductive gel. One main advantage of low impedance is lower noise in electrode potentials due to lower resistance (Kappenman & Luck, 2010).

Another advantage of low impedance is lower contamination of the EEG signal by skin potentials. Skin potentials arise when a participant sweats, which changes the skin's impedance and the voltage on the scalp. Skin potentials can also arise if an electrode moves slightly on the scalp: the voltage changes if the skin underneath the electrode has different impedance to the previous electrode position. Skin potentials are often large and are a source of low-frequency noise in EEG recordings. Therefore, it is desirable to reduce their occurrence. Ensuring low impedances for each electrode at the start of an experiment generally results in smaller changes in impedance due to sweating (Picton & Hillyard, 1972).

2.3. EEG processing

Pre-processing refers to a variety of techniques for 'cleaning up' raw EEG data. It is common practice to use pre-processing techniques before drawing statistical comparisons. A range of considerations were taken into account when processing the EEG data recorded in the experiments reported in this thesis. The principles and main issues are described in this section; the details are given in the methods sections of subsequent chapters.

2.3.1. Choice of reference site

Ideally, the reference electrode should be independent from the recording electrode because EEG waveforms depend on activity at a reference electrode in addition to activity at the recording electrode. However, it is difficult to locate an electrically neutral reference site. The following two sub-sections will evaluate the theoretical justifications for the two most commonly-used reference sites: the mastoid reference and the average reference.

2.3.1.1. Mastoid reference

Historically, one of the most common reference positions was the mastoids. The mastoids are conical protrusions of the skull located just behind the ears (Figure 2.5). The mathematically-linked mastoid reference takes the average of recordings gained separately from the left and right mastoids. The rationale for selecting the mastoids as a reference is that they are located on the head so they pick up some of the noise that is present in the recording electrodes, but it is assumed that they do not pick up signals of interest from the brain.

The assumption that the mastoids do not pick up signals of interest from the brain, however, has not been confirmed. For example, Srinivasan, Nunez, and Silberstein (1998) simulated 4240 dipole sources. They simulated 500 random source distributions, from which they estimated scalp potentials for 111 electrodes, and calculated coherence values (i.e. squared correlation coefficients) between all possible pairs of electrodes. Using a linked mastoids reference, approximately half of the coherence values differed by more than 0.1 from reference-free coherence estimates. For electrodes less than 15 cm apart, the mastoids reference conflated coherence values, whereas for electrodes separated by longer distances, coherence values were smaller than reference-free estimates. These results suggest that the mastoid reference may artificially correlate data from some electrodes, which can lead to inaccurate coherence estimates. These artificial correlations also affect the distribution of scalp maps (Nunez & Srinivasan, 2006). When Srinivasan *et al.* (1998) simulated the average reference, however, all of the coherence values were within 0.1 of the reference-free estimates. This result suggests that the average reference may be a better approximation of reference-free recordings than the mastoid reference.

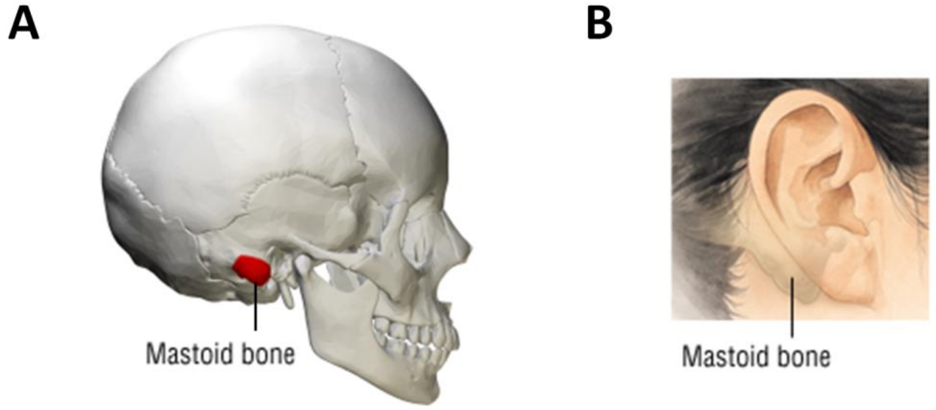


Figure 2.5. Diagram showing position of right mastoid on the skull (**A**; adapted from http://en.wikipedia.org/wiki/Mastoid_process [retrieved 14/08/2014]) and on the skin (**B**; adapted from http://www.drugs.com/health-guide/chronic-otitis-media-cholesteatoma-and-mastoiditis.html [retrieved 14/08/2014]).

2.3.1.2. Average reference

The average reference uses the average voltage at all of the recording electrodes as a reference. This method has the advantage that it less biased by potentials at single recording sites than other references, such as the mastoids reference. However, some researchers have argued that the average reference can lead to misinterpretations when it is computed from a small number of electrodes that do not adequately cover the head (Desmedt, Chalklin, & Tomberg, 1990). Although, other researchers have argued that the average reference approximates reference-free recordings when recording from dense electrode arrays (e.g. Bertrand, Perrin, & Pernier, 1985; Dien, 1998), which is consistent with the results of Srinivasan *et al.* (1998).

Performing source reconstruction with EEG data (discussed later in this chapter) requires the computation of a mathematical formula that is equivalent to first performing the average reference. Therefore, if a researcher wishes to perform source reconstruction, it would be consistent to use the average reference when interpreting scalp-recorded EEG data (Handy, 2009; Michel et al., 2004). For that reason, the experiments reported in this thesis were conducted using the average reference method.

2.3.2. Filtering

In EEG research, the data of interest is often contained within a specific frequency band. For event-related potentials (ERPs), the relevant part of the waveform typically lies between 0.01 and 30 Hertz (Hz; Luck, 2005). Filtering the raw EEG data can be useful for removing noise that is likely to occur at specific frequencies. Low-pass filters retain frequencies lower than the specified value and high-pass filters retain frequencies higher than the specified value. Bandpass filters retain frequencies within a certain range.

High-pass filters are useful for removing direct current (DC) offset and artifacts arising from skin potentials, which are typically low-frequency (< 0.1 Hz). Low-pass filters can be used to remove frequencies containing line noise caused by electrical equipment, such as monitors and cables (~50 Hz). Nevertheless, filtering should be applied with caution, since unnecessary filtering can distort EEG waveforms. When considering the appropriate filters to apply, the Nyquist theorem should be taken into account to avoid distortion in the frequencies of interest. Luck (2005) points out several potential consequences of inappropriately filtering ERP waveforms. Filtering may change the apparent timing of ERPs, induce artificial oscillations, and make monophasic waves appear multiphasic. Therefore, it is important to consider the possible consequences of filtering and check that the filtering technique has not altered key aspects of the waveform, particularly aspects that contribute to subsequent statistical comparisons.

2.3.3. Creating epochs

When studying evoked responses, it is necessary to create a time window around the event of interest, which in cognitive neuroscience experiments is typically the time of stimulus onset or the time at which a participant makes a behavioural response. When selecting an epoch, it is important to also include a 'neutral' time window to act as a baseline for each epoch, when it is assumed that the processes that the researcher wishes to measure have not yet begun.

2.3.4. Artifact removal

Artifacts are not always limited to a specific frequency band. Even if they are, the frequencies containing artifacts might overlap with the frequencies of interest, meaning that it would not be desirable, or perhaps possible, to filter them out without distorting the waveform substantially. Many artifacts result from muscle movements—for example, eye blinks and eye movements. These artifacts are problematic because they typically have much greater amplitudes than scalp-recorded neural activity.

One option is to remove all epochs suspected to contain eye blinks, which can be detected by high-amplitude peaks. This method is commonly referred to as artifact rejection. However, removing all eye blinks often means rejecting a large proportion of epochs, which decreases the signal-to-noise ratio (SNR) of the resulting averaged data.

Artifact correction refers to techniques designed to isolate artifactual components and remove them from the data. Instead of removing all epochs containing artifacts, the artifactual component is removed from every epoch. Independent component analysis (ICA) is the most commonly-used method for detecting artifacts. ICA aims to decompose a dataset into separate independent signals. These signals are assumed to be mixed linearly in the dataset. The technique outputs a set of components that are statistically independent[3] from each other (Neuper & Klimesch, 2006). Artifact correction refers to the method of inspecting the waveforms associated with the independent components, identifying one or more waveforms that are suspected to reflect artifactual activity, and removing these waveforms from the EEG data. The modified dataset consists of a linear mixture of the remaining ICA components, which are assumed to reflect EEG activity generated by brain potentials.

ICA is commonly used to correct for eye blinks, which have a stereotyped scalp distribution. Eye blinks generate electrical activity of opposite polarity at sites above and below the eyes, whereas brain potentials typically have similar voltages above and below the

---

[3] The fastICA algorithm, which is applied to the EEG data in this thesis, defines statistical independence by maximising non-Gaussianity (Hyvärinen & Oja, 2000).

eyes (Luck, 2005). By plotting the scalp distribution of each ICA component, it is possible to infer which component(s) are most likely associated with eye blinks. However, as with all EEG processing techniques, ICA should be used with caution. If a participant blinks in a non-random pattern, for example at stimulus onset, then aspects of the evoked response might be correlated with eye blinks; thus, one ICA component could contain a mixture of eye blink signals and brain potentials. Consequently, removing that ICA component might also remove aspects of EEG activity that arise from neural activity. On the other hand, removing an ICA component might undercorrect for eye-blink artifacts (Keil, Müller, Ihssen, & Weisz, 2012), meaning that artifacts still remain in the data. Nonetheless, if most of the artifactual activity has been removed, then any remaining noise should cancel out with averaging. Before conducting ICA, it is important to check that eye blinks are not time-locked to parts of the epoch—for statistical comparisons, time-locked artifacts might be particularly problematic if two experimental conditions differ in the extent to which they evoke artifacts. As a result, ICA over- or under-corrections might introduce differences between the two conditions. For this reason, epochs should be visually-inspected for artifacts before proceeding with further analyses.

Apart from these cautions, several studies have reported that ICA is able to successfully remove eye-blink and eye movement artifacts (e.g. Hoffmann & Falkenstein, 2008; Jung, Makeig, Humphries, et al., 2000; Jung, Makeig, Westerfield, et al., 2000). One way in which the quality of ICA results can be improved is by using a larger number of time points (Neuper & Klimesch, 2006), such as by using a high sampling rate. Hyvärinen and Oja (2000) also speculate that the quality of ICA can be improved by bandpass filtering the data. For the EEG experiments reported in this thesis, both of these techniques were employed.

Artifact correction can be particularly useful when analysing data from populations whose data is likely to be heavily contaminated with eye blinks, such as children (Handy, 2009). In these cases, fewer trials may be able to be gained than with adult populations, which means that the SNR will be disadvantageous, even preceding artifact analyses. Artifact rejection can decrease the SNR further. However, if artifact correction rather than artifact rejection is applied, then it is possible to remove eye-blinks while largely preserving the SNR of participant averages.

2.3.5. Averaging

Evoked responses can be analysed by averaging across all trials for an individual participant. The resulting waveforms are referred to as ERPs. By definition, averaging cancels random noise. Therefore, if large numbers of epochs are averaged, random noise will be reduced in the average waveform. The amount of random noise present in an average waveform decreases as a function of the square root of the number of trials in the average (beim Graben, 2001; Niedermeyer & Lopes da Silva, 1999). However, signal-correlated non-random components remain in the average waveform. These components are assumed to reflect the signal of interest in an experimental task (assuming that the previous pre-processing steps have been applied appropriately). As a result, it is desirable to collect a large number of trials in EEG experiments since this leads to better SNR for average waveforms.

Some researchers also average across participants to create 'grand average' waveforms. The problem with this type of averaging is that it masks individual variability between participants. In a grand average waveform, it is easy to see similarities between participants, however, this waveform might not accurately reflect the pattern of results contained in any of the data from individual participants (Luck, 2005). For example, if the data from half of the participants have a monophasic peak at 50 ms and the other half at 100 ms, then this might appear in the grand average as a biphasic peak with lower amplitude. The grand average has the potential to be incorrectly interpreted if the individual waveforms are ignored.

2.4. Methods of analysis

This section will discuss some methods of analysis for making statistical comparisons among evoked responses. One common method for statistically comparing ERPs is to identify and analyse ERP components. However, I argue that there are several disadvantages to this technique. I then argue that Cluster-based Permutation Analysis is a preferable alternative. This is the technique that was used to analyse the scalp-recorded EEG data in this thesis. The second half of this section will consider one way in which inferences can be drawn about the neural generators of scalp-recorded activity. An increasingly-common approach is to use an inverse model to reconstruct the sources at many voxels in the brain.

2.4.1. ERP components

2.4.1.1. Definition and uses

Luck (2005) defines an ERP component as a single cognitive operation that influences scalp-recorded voltages. Like most current cognitive neuroscience methods, EEG does not provide access to individual operations, but a combination of operations that are involved in a task. ERP research often aims to isolate a single component from an ERP waveform and investigate changes in the amplitude and/or latency of that component under different experimental conditions. ERP components are generally given names that indicate their polarity (positive = P; negative = N) and their timing or position within the waveform (e.g. P300, which peaks approximately 300 ms after stimulus onset).

ERP amplitudes are typically thought to indicate the strength of the neural response (Luck, 2005). A component's latency (defined as the time taken to reach peak amplitude) is thought to measure the amount of time from stimulus onset to the brain response. Since electricity travels almost at the speed of light, scalp-recorded activity reflects activity that is happening in the brain practically at the time of recording (Luck, 2005). Researchers typically conduct a *t*-test or ANOVA on amplitude or latency measures in order to test for a significant difference between two experimental conditions.

2.4.1.2. Previous findings

Decades of research have contributed to the identification of a large number of ERP components that are observed consistently across experiments. For example, the P1-N1-P2 complex is thought to reflect stimulus processing in sensory cortices. These components occur within 150 ms after stimulus onset and are affected by properties of the stimulus, such as frequency, duration, intensity, and location (McEvoy, Picton, Champagne, Kellett, & Kelly, 1990; Ostroff, McDonald, Schneider, & Alain, 2003; Taub & Raab, 2005; Woods et al., 1993). Longer-latency components are thought to represent cognitive processing. For example, the P300 has been associated with exogenous attention to a salient stimulus (Friedman, Cycowicz, & Gaeta, 2001; Polich, 2009; Squires, Squires, & Hillyard, 1975).

2.4.1.3. Limitations of ERP component analysis

Isolating ERP components can be challenging because several components are typically present in a single ERP waveform. These components can be temporally overlapping and can span similar groups of scalp locations. Therefore, it can be difficult to tease apart the explanations that an identified ERP component reflects a single cognitive operation compared to several overlapping components. Previous research has identified components that were once thought to reflect a single cognitive process, that in fact are now believed to reflect several cognitive processes. For example, Näätänen, Gaillard, and Mäntysalo (1978) proposed that the modulation of N1 previously described by Hillyard *et al.* (1973) could actually be explained by modulation of a negative component that overlapped in time with the classic N1 component. Using a longer inter-stimulus interval than Hillyard *et al.* allowed Näätänen *et al.* to tease apart the two components in time and scalp distribution.

The distinction between a single component and an overlapping set of components becomes important when drawing conclusions about the reasons why two waveforms differ. An apparent reduction in amplitude could be the result of several processes: (1) a reduction in the response of one brain region that is producing a single component; (2) the recruitment

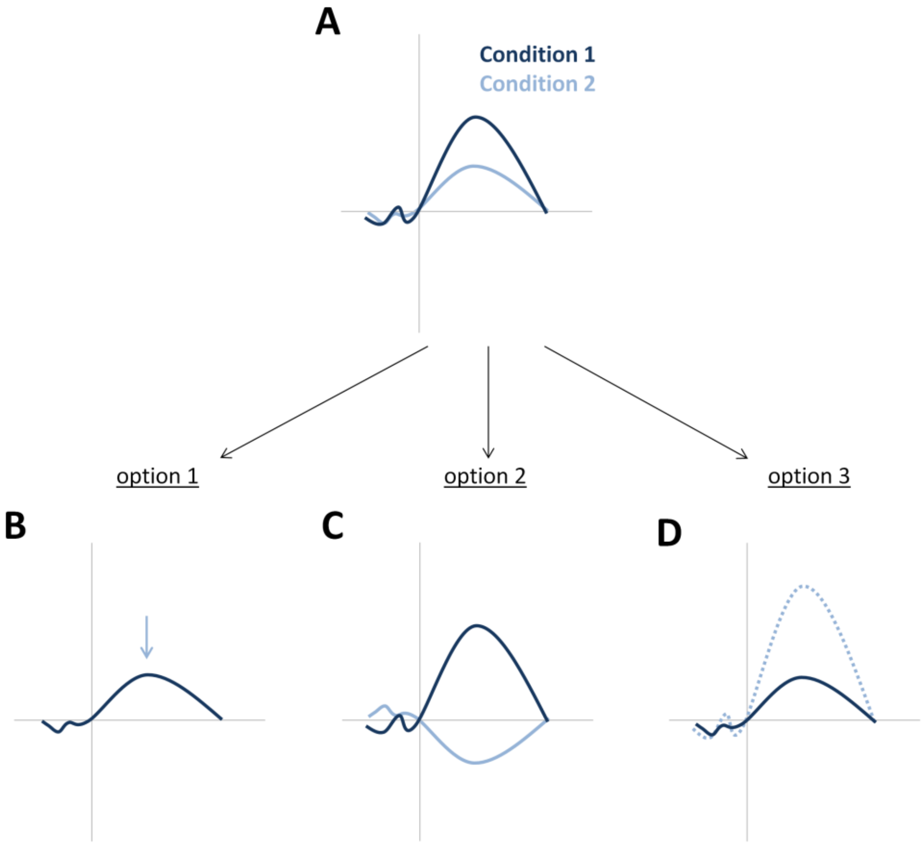


Figure 2.6. Schematic showing three possibilities for the change in underlying components that might result in an apparent decrease in amplitude in the waveforms from one experimental condition (Condition 1) to another (Condition 2; **A**). (**B**) One option is a single underlying component in Condition 1 that decreases in amplitude in Condition 2. (**C**) A second option is the addition of a component with negative polarity in Condition 2. (**D**) A third option is the presence of two overlapping positive components in Condition 1, of which the higher-amplitude component disappears (dotted line) in Condition 2.

of additional brain regions to perform different operations at a similar time but with opposite polarity on the scalp; or (3) elimination of one positive scalp component, revealing a lower-amplitude positive component underneath (Figure 2.6). Similarly, a change in multiple peaks in an ERP waveform might reflect a change in one sustained component rather than multiple separate modulations (Figure 2.7). The results of Hillyard *et al.* (1973) and Näätänen *et al.* (1978) highlight how conclusions about the effects of an experimental manipulation have the potential to be misinterpreted when two or more components overlap in time. The effects of an experimental manipulation may be misattributed to a component that is unaffected by the experimental manipulation, which has the potential lead to erroneous conclusions about the cognitive processes that underlie the experimental manipulation.

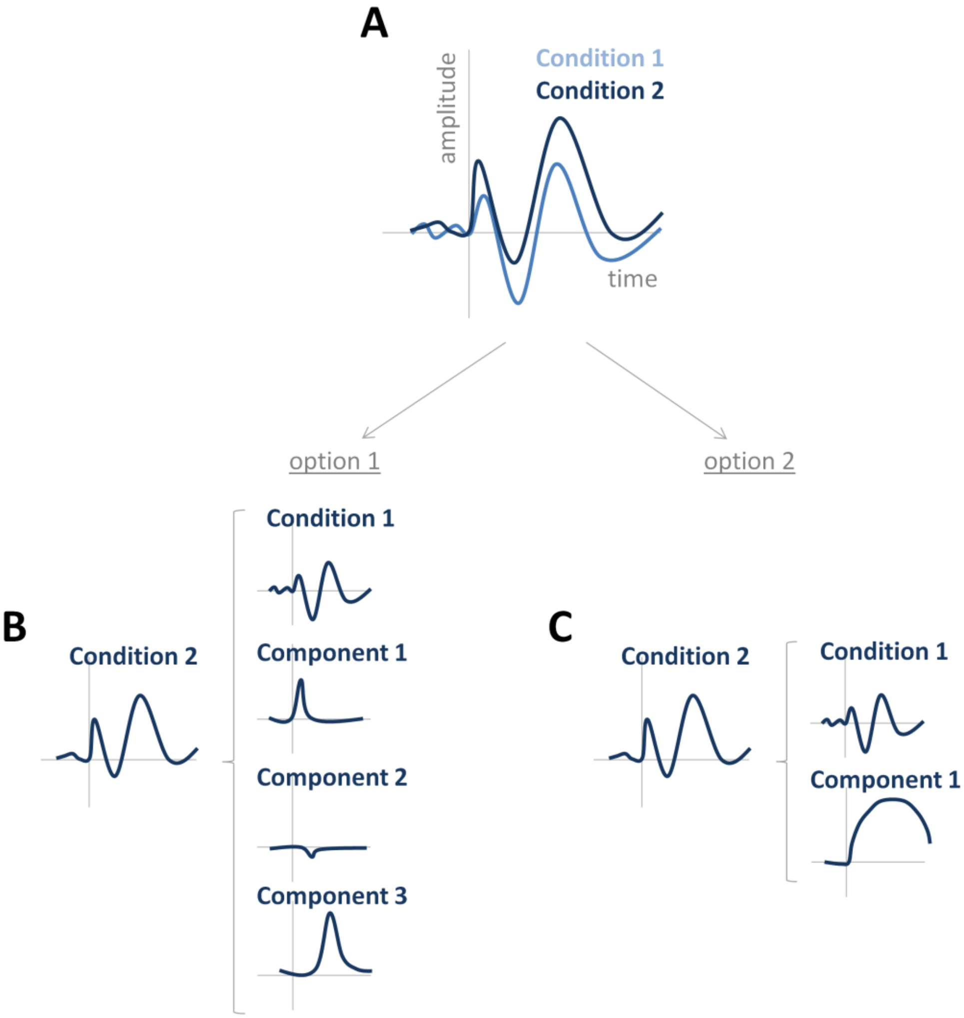


Figure 2.7. Schematic showing two possibilities for a difference in the amplitude of multiple peaks between two conditions. (**A**) A schematic of hypothetical waveforms observed in two conditions. (**B**) One option is that Condition 2 reflects the waveform of Condition 1 with larger positive peak amplitudes and smaller negative amplitudes that affect each peak separately. (**C**) A second option is the addition of a sustained component that affects all peaks simultaneously.

The overlapping nature of components also presents a second problem: there is not an obvious relationship between the peak of an averaged waveform and the peak of an ERP component (Luck, 2005). If part of the waveform is a mixture of different underlying components, then the peak of a component does not necessarily occur at the same time as the peak in the waveform, which could lead to erroneous estimations of latency for a particular cognitive process (Figure 2.8). In a similar manner, a change in component amplitude can result in an apparent change in latency and *vice versa* (Luck, 2005; Figure 2.9). Since

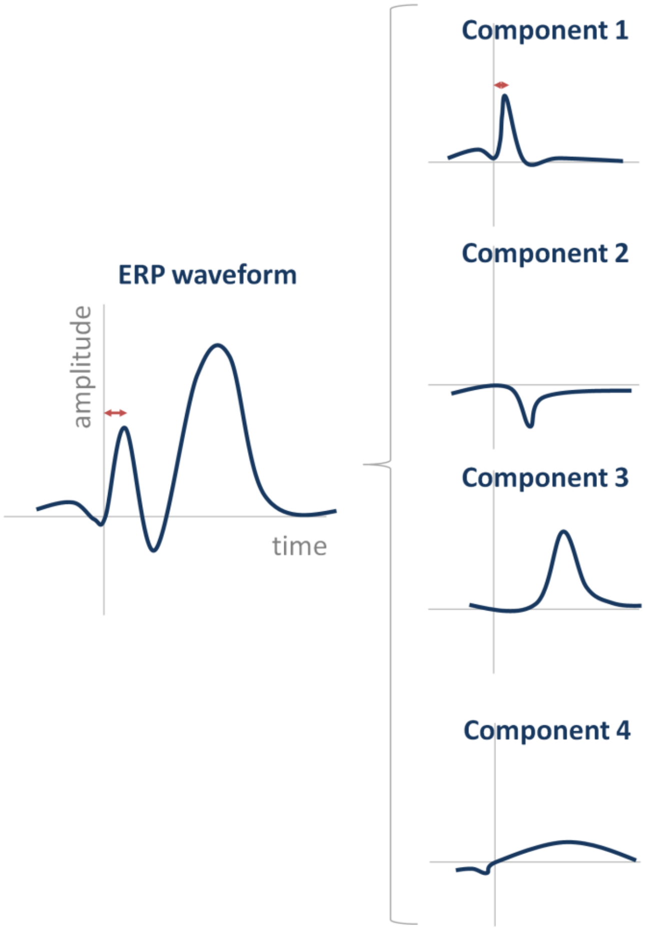


Figure 2.8. Schematic showing a hypothetical ERP waveform (left) that is comprised of four major components (right). As a result of Component 4, which is sustained over the time window, the latency of peaks in the ERP waveform do not match the latency of peaks of individual components. Arrows display the difference in latency between the first positive peak in the ERP waveform and the peak of Component 1.

waveform peaks and components are not the same, it does not make sense to focus only on waveform peaks at the exclusion of other parts of the waveform.

One method by which parts of the waveform are able to be segregated is by measuring the relative timing of sinusoidal components of different frequencies. For example, phase delay is a measure of the time delay of the phases and group delay is a measure of the time delay of the amplitude envelopes. However, although these methods are able to decompose a waveform into a series of different components, few researchers employ these methods when analysing ERP waveforms.

Another challenge for ERP researchers is deciding whether the component present in one experiment is the same as a component that has been identified in previous experiments. In particular, it should not be assumed that components with similar timing and polarity reflect the same underlying sources. One classic example is provided by the early sensory components: the P1, N1, and P2 responses are observed for both visual and acoustical stimuli. However, just because the labelling is the same does not mean that these components are functionally similar across modalities or that they are generated by similar regions of cortex.

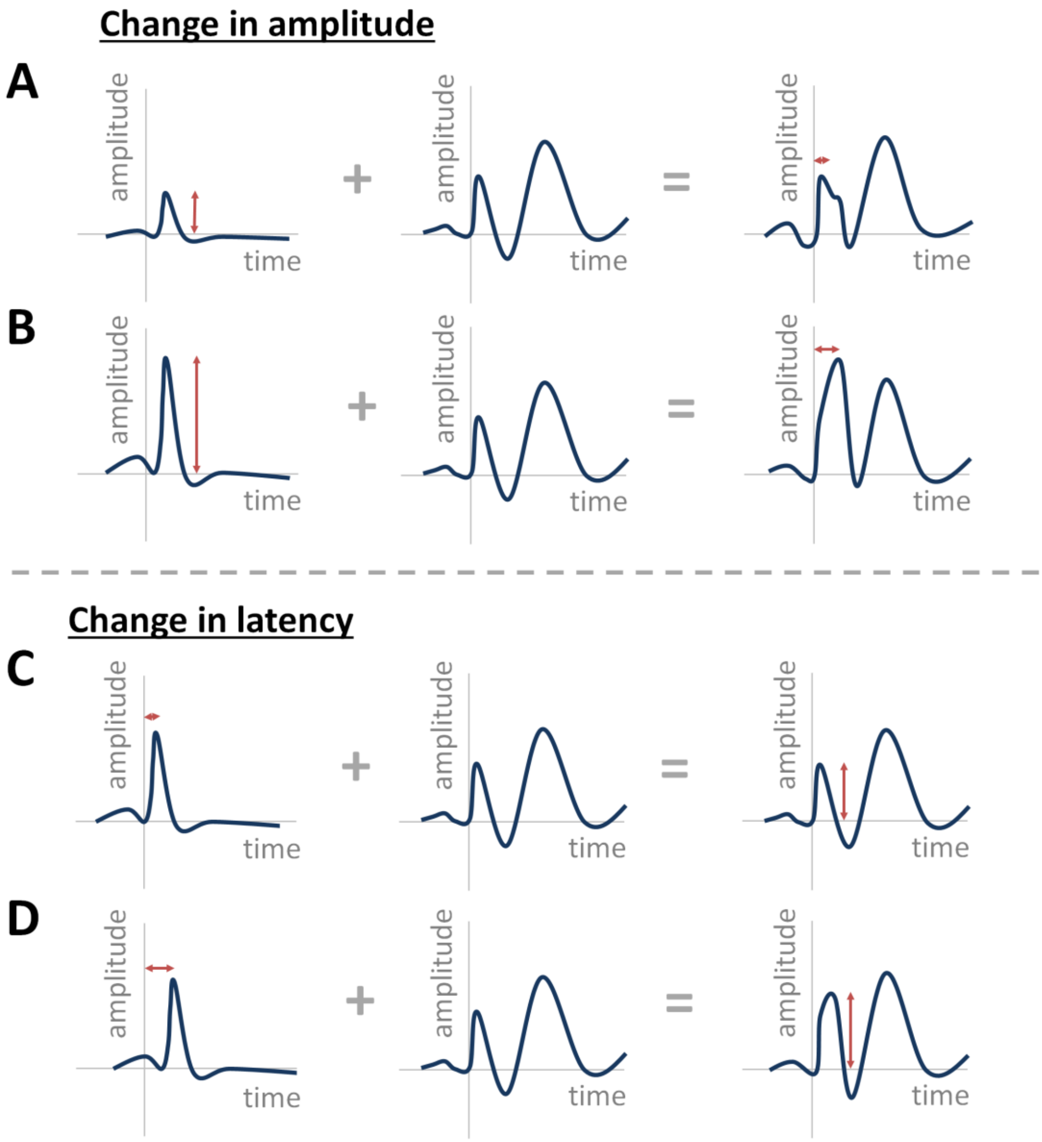


Figure 2.9. Schematic showing how adding different components to the same ERP waveform can have different, and possibly unexpected, effects on the resulting waveform (right panel). (**A**)-(**B**) show how a change in component amplitude can affect the latency of the resulting waveform. (**C**)-(**D**) show how a change in component latency can affect the amplitude of the resulting waveform.

Therefore, similar labelling does not mean that two components reflect similar underlying processes. A component's topography (i.e. its distribution across the scalp) reflects properties of the neural generator(s) in the brain that produced it and can be used to inform decisions about whether two components are the same.

Finally, the interpretation of the underlying processes of ERP components mainly depends on knowledge of the components based on previous literature. One consequence of this approach is that the conclusions drawn from the results are only as strong as previous experiments that have made inferences about the processes underling these components. Secondly, it relies on confidence that the components from the different experiments are, indeed, the same functional component. Not all experiments identifying ERP components rely on the identification of a specific component, but many do.

Given the aforementioned limitations in identifying ERP components, this method may not be suitable for all experiments. For example, when the cognitive process of interest has not been widely studied in the EEG literature or when there are no prior expectations about the electrodes and time points at which experimental conditions are expected to differ.

2.4.2. Cluster-based Permutation Analyses

Cluster-based Permutation Analyses can search through the data space and identify time points where two conditions differ, irrespective of whether those time points lie at the peak of the average waveform or not. Advances in EEG technology mean that researchers commonly record from 64 or more electrodes, thereby achieving high spatial resolution on the scalp. This poses a problem for traditional analyses, where the researcher typically selects or averages across electrodes of interest. The alternative approach of clustering in the spatial dimension allows the researcher to enter every electrode into the analysis, while avoiding multiple comparisons. By exploiting these principles, Spatio-temporal Cluster-based Permutation Analysis (Maris & Oostenveld, 2007) identifies clusters of electrodes, grouped over space and time, where activity differs systematically between two conditions.

The analysis entails five steps. First, the mean amplitudes of the EEG signal in the two conditions of interest are compared at every space-by-time point (i.e. at each time point at each electrode), for example, in a paired-samples *t*-test[4]. This method identifies space-by-time points that show significant differences in amplitude between the two conditions at a predefined level[5]. The level specified in this thesis is the $p < 0.05$ (uncorrected) level. The space-by-time points that exceed the specified level are retained for further analysis. At the second step, the retained points are grouped into “clusters”, provided that they occupy neighbouring time points at the same space point, or occupy neighbouring space points at the same time point. Thus, clusters can span both spatial and temporal dimensions, but the points that define them are always connected both spatially and temporally.

At the third step, a *cluster statistic* is calculated for each cluster as the sum of the *t*-values of the points included in the cluster. Fourth, to create a null distribution against

---

[4] The paired-samples *t*-test is the method applied to the EEG data in this thesis. Although, the Cluster-based Permutation Analysis has the potential to be applied to the results of any test statistic.

[5] This level does not affect the false alarm rate of the statistical test. However, it does affect its sensitivity (Maris & Oostenveld, 2007). For example, weak effects that are sustained over long durations will not be detected if a high threshold is specified.

which to test the cluster statistic, the original data are repeatedly permuted. For each permutation, the average data for the two conditions are swapped at every space-by-time point for a randomly selected subset of the participants (where the number of participants whose data was swapped can take any value between 1 and the total number of participants). The rationale behind permutation is that, under the null hypothesis, allocating the data into separate conditions is arbitrary, so swapping the labels should have no effect on the size of the clusters. Permuting the data creates a new set of clusters and, thus, a new set of cluster statistics. The maximum number of possible permutations is equal to 2^N, where N is the total number of participants. However, with large numbers of participants, the maximum may be impractically large and is rarely performed. Based on the results of simulations and applications to experimental data, Marozzi (2004) suggests that 5000 permutations is sufficient to provide reliable *p*-value estimates when alpha is equal to 0.05. The maximum cluster statistics from the selected number of permutations[6] are compiled to form a non-parametric probability distribution (referred to as the 'null' distribution). At the final step, the cluster statistics from the observed data are compared against this distribution in a two-tailed test. If the maximum number of permutations was not applied, then the *p*-values of the observed clusters are approximated by the Monte Carlo estimate, as the proportion of the distribution that is larger than or equal to the observed cluster statistic.

Spatio-temporal Cluster-based Permutation Analysis resolves the multiple-comparisons problem because the difference between conditions for each cluster is evaluated by a single test statistic—the cluster statistic—that encompasses the entire spatio-temporal array of the cluster. Multiple comparisons at each space-by-time point are replaced by a single comparison—it is the *clusters* of space-by-time points that are compared against the null distribution, rather than individual space-by-time points. If a cluster is larger than expected from a comparison with the null distribution, then the cluster is considered significant. The comparison takes account both of the magnitude of the *t*-values at each point in the cluster and of the extent of the cluster over space and time. Clusters can reach the criterion for significance either if they display modest *t*-values over a large number of neighbouring space-by-time points, or if they display large *t*-values over a smaller group of neighbouring space-by-time points.

Other advantages of permutation analyses are that they do not require a specific underlying distribution, they incorporate no *a priori* assumptions about when or where an effect is likely to occur, and they can cope with large numbers of electrodes and time points

---

[6] The number of permutations used in the experiments in this thesis was 10000, which allowed a more precise estimate of the *p*-value that resulted from the permutation analysis than with 5000 permutations.

with no increase in the proportion of Type-I errors. For these reasons, Spatio-temporal Cluster-based Permutation Analyses were used to identify differences in ERP waveforms throughout this thesis.

2.4.3. Source reconstruction

Source reconstruction allows researchers to estimate the neural sources that underlie scalp-recorded activity. However, source reconstruction is not a straightforward computation.

2.4.3.1. Forward model

In order to map scalp-recorded activity into source space, researchers must first develop a model that specifies how neural activity produces scalp-recorded activity. The parameters of this model are referred to as the ‘forward model’. The aim of the forward model is to estimate the geometric and conductive properties of the head. These properties are expressed in the lead field matrix, which specifies the relationship between activity at each possible source location and the resulting amplitude at each electrode on the scalp. The lead field matrix is multiplied with a matrix of source estimates to produce a forward solution (Michel et al., 2004). The ‘error’ between predicted and measured scalp potentials forms the basis for localising scalp-recorded potentials.

The simplest forward models are spherical models, which estimate the head using between one and four concentric shells. The four-sphere model specifies different conductivities for the brain, skull, scalp, and cerebrospinal fluid, since these are where the most prominent differences in conductivity occur. Each shell, however, is assumed to be homogenous, which is an oversimplification of conductivity in the human head (Keil et al., 2013).

As a result of greater computational processing ability, realistic head models are becoming increasingly popular. One example is the boundary element method (BEM). It models the brain as a triangular mesh with different conductivities, which is combined with a structural MRI image to restrict the source space to likely EEG generators (Fuchs, Wagner, & Kastner, 2001). The use of complex head models is thought to provide a more accurate estimate than simple spherical models (Cuffin, 1993, 1996; Menninghaus, Lütkenhöner, & Gonzalez, 1994). For example, Cuffin (1996) recorded EEG at 16 electrodes from three epilepsy patients who were implanted with depth electrodes at known locations. Realistic and spherical head models were applied to attempt to reconstruct the depth electrode source. They found that, provided the SNR was greater than 50 (calculated as half of the largest peak-to-peak amplitude in the experimentally recorded EEG divided by the root-mean-square during the pre-stimulus baseline), the realistic head model was more accurate than the spherical model at predicting the location of the depth electrode. The BEM model uses geometrical information from structural MRI scans to inform the forward model. In many experiments, the structural MRI is that of a 'template' MRI scan, which consists of an average over many participants or a representative single-subject scan. However, when individual structural MRI scans are available for each participant in an EEG experiment, these scans can be used to produce an individual head model for each participant, based on the geometrical information contained in each participant's MRI scan. Michel (2004) assumes that the advantage of using individual MRI scans over simpler models is that it restricts the source space to locations in which EEG sources can arise for each individual, such as grey matter and some sub-cortical structures.

2.4.3.2. Inverse problem

The 'inverse problem' arises because recorded amplitudes from the scalp do not directly reflect the underlying neural generators in the brain. This is an underdetermined problem, meaning that the number of electrodes from which amplitudes are measured is insufficient to uniquely identify the configuration of neural generators. This problem occurs because the folding of the cortex means that activity in one source can be cancelled by activity in another, leading to no detectable difference on the scalp (Luck, 2005). Inverse models must specify additional constraints in order to produce a unique solution (Hämäläinen & Ilmoniemi, 1984; Helmholtz, 1853). One strategy is to constrain the solution space to cortex. This reduces the number of possible solutions, although the number of possible solutions still remains large.

The minimum norm approach (Hämäläinen & Ilmoniemi, 1984) provides an additional constraint. It minimises the source variances. The logic of this constraint is that cancellation of sources means that estimated magnitudes could be very large, but cancel each other out. Magnitudes that are very large are biologically implausible (Luck, 2005). One consequence of the minimum norm approach is that the solution is biased towards sources that are close to the cortical surface, because deeper sources require greater magnitudes to reach the scalp. To overcome this bias, depth-weighting strategies have been applied to the minimum norm approach (Lin et al., 2006). The minimum norm always provides a unique solution because only one solution perfectly fits the scalp-recorded data and also produces the smallest overall source magnitude.

2.4.3.3. Correcting for multiple comparisons

Once an inverse model has been applied to the data, the next step is to statistically analyse the source space estimates. This is typically carried out using an ANOVA or *t*-test at each voxel of reconstructed activity. However, if the researcher wishes to search the entire source space (i.e. with no *a priori* assumptions about the brain regions active; often referred to as a 'whole brain analysis'), then a correction for multiple comparisons must be applied. The correction should control for the occurrence of Type-II errors. However, corrections also reduce the statistical power, which can become a problem for EEG source analysis if reconstructed source estimates for individual participants are noisy.

2.5. Application of EEG to the current project

One of the main advantages of cognitive neuroscience methods is that they can measure neural activity during a task without requiring a behavioural response. Analysing brain activity and manipulating the direction of attention allows researchers to investigate the timing of different attentional processes, the functional location of these processes in the brain, and the degree to which the neural response reflects aspects of attended and unattended stimuli (e.g. Hill & Miller, 2010; Srinivasan et al., 2009; Woldorff et al., 1993).

The high temporal resolution of EEG is particularly relevant for the aims of this thesis. In previous experiments, preparatory and selective attention occurred within a single trial of a multi-talker listening task (Hill & Miller, 2010; Lee et al., 2013). To investigate preparatory and selective attention in normally-hearing listeners, it is important to separate brain activity that occurs before a talker starts to speak from activity that occurs while one or more talkers are speaking. EEG allows processes that occur at different time points within a trial to be separated because EEG measurements occur, for practical purposes, at the same time that electrical activity occurred in the brain. This is one advantage of EEG over haemodynamic measures, such as fMRI, where the BOLD response might take seconds to manifest. While one previous experiment has attempted to localise the brain network during preparatory attention using fMRI (Hill & Miller, 2010), the time course of this process is currently unknown. Therefore, when measuring the time course of preparatory attention in the brain, EEG is an important complement to fMRI.

Although EEG is not ideal for localising the sources of neural activity, Michel et al. (2004) proposed that recent technological advances have vastly improved its spatial resolution. High-density electrode recordings, realistic forward models, and modern inverse solutions all contribute to this improvement. Previous studies have been able to identify some correspondence between source reconstruction of EEG activity and fMRI results during equivalent tasks (Lachaux et al., 2007; Rusiniak et al., 2013). For example, Rusiniak *et al.* (2013) recorded simultaneous fMRI from eleven children. During an oddball task, the P300 component of the EEG response was localised to similar parietal areas that showed increased BOLD activity using fMRI.

While MEG has high temporal resolution, the inverse solution is underdetermined, similar to EEG. EEG and MEG are preferentially sensitive to different source orientations. Therefore, they provide non-redundant information and the most comprehensive information about source location arises from combining the two techniques. A previous experiment investigated preparatory attention using MEG (Lee et al., 2013). Therefore, gathering additional information from EEG will improve knowledge of this topic.

In contrast to other cognitive neuroimaging techniques (e.g. MEG and fMRI), presenting acoustical stimuli to participants through loudspeakers is not a problem for EEG recordings. This allows for more realistic listening environments, where different sounds are presented from different locations in space.

2.5.1. Child EEG

Another advantage of EEG is that it is well-suited to recording brain activity from children. Firstly, the child's head does not need to be restrained. Secondly, active EEG systems can at least partially correct for head movements, which tend to occur more frequently in children than adults. In addition, EEG recordings can be made while a hearing-impaired child is wearing hearing aids.

One consideration with child EEG is increased noise. Children generally cannot complete as many trials in a session as adults, which may lead to lower SNR in the average waveforms. Greater proportions of trials containing artifacts may also lead to noisier data with greater variability (Coch & Gullick, 2011; Luck, 2005). If all trials containing artifacts are removed from child EEG data (i.e. artifact rejection), then this method might further decrease the SNR in the average waveforms. Thus, artifact correction (Section 2.3.4) is a particularly useful technique for EEG data obtained from children. Correcting for artifacts, rather than discarding all trials containing artifacts, is a method that aims to remove artifactual components without decreasing the SNR of the average waveforms. In this thesis, the criterion level for artifact rejection was varied to maintain a similar proportion of rejected trials for children and adults. This method removed trials that contained the highest-amplitude artifacts in each child's EEG data, but did not lead to the rejection of a high proportion of all trials. Since it was estimated that artifacts remained in the data, ICA was applied to the EEG data from every child to correct for remaining eye-blink artifacts, while for adults, ICA was only applied to participants who demonstrated trials that contained eye blinks.

2.6. Summary and conclusions

EEG is well-suited to the aims of this thesis. Firstly, it measures electrical activity at the scalp that is a direct measure of underlying neural activity. Secondly, high temporal resolution makes EEG ideal for exploring the time course of preparatory attention. Also, sounds can be presented in the sound field, which more easily approximates everyday multi-talker listening situations than simulating spatial locations through headphones. Finally, EEG is well-suited to measuring brain activity from adults, from children, and from hearing-impaired children when listening with and without their hearing aids.

This thesis employed the pre-processing and analysis methods discussed in this chapter. Spatio-temporal Cluster-based Permutation Analysis was used to analyse ERPs, since there were no *a priori* assumptions about the electrodes or latencies at which significant differences were expected to occur. In one experiment where individual structural MRI scans were available, this thesis also estimated the neural generators that contributed to significant differences recorded at the scalp.

The data gained from EEG recordings complements the results of previous experiments that have employed MEG and fMRI. One aim of using EEG throughout this thesis was to gain more information about the time course of brain activity for preparatory and selective attention during multi-talker listening than has been reported in previous experiments.

Chapter 3

EEG Activity during Two-talker Listening

The aim of the three experiments presented in this chapter was to measure the temporal dynamics of brain activity during two-talker listening—in young adults and children aged 7–13 years. Previous experiments with adults have shown improved speech intelligibility from knowing the spatial location (Ericson et al. 2004; Kidd et al. 2005; Best et al., 2007; Best et al., 2009) and the identity (Kitterick, Bailey, & Summerfield, 2010) of a target talker before he or she begins to speak. Although these behavioural advantages have been observed consistently, the neural processes that underpin them are not fully understood.

Two previous experiments have studied brain activity evoked by preparatory attention during multi-talker listening using functional magnetic resonance imaging (fMRI; Hill & Miller, 2010) and magneto-encephalography (MEG; Lee et al., 2013). Hill and Miller (2010) presented three simultaneous talkers, which differed in simulated spatial location and average fundamental frequency (F0). Before the talkers began, a visual cue indicated either the location (left/right/centre) or the F0 (high/low/middle) of the target talker. The visual cue evoked activity in a left-hemisphere fronto-parietal network when participants were cued to location and F0. However, the detailed pattern of activity within the network depended on whether participants were preparing to select the upcoming target talker based on location or F0. Thus, the results provide evidence for both domain-general and cue-specific brain activity. Lee *et al.* (2013) used a similar task, but presented two simultaneous spoken digits on each trial. Lee *et al.* found greater preparatory activity in the left dorsal precentral sulcus and gyrus during attend-location trials and in the left posterior superior temporal sulcus during attend-F0 trials. These results, like Hill and Miller's, demonstrate cue-specific brain activity during preparatory attention.

These experiments aimed to build upon the results of Hill and Miller (2010) and Lee *et al.* (2013) in two respects. First, neither Hill and Miller or Lee *et al.* addressed the question of how soon attentional preparation is manifest in neural activity. Second, a possible shortcoming of the experiments of Hill and Miller and Lee *et al.* is that differences in the feature to be used for selection (i.e. location or F0) were confounded with differences in the visual cues (i.e. differences in chevron orientation). The latter was a key issue in the design of the current experiments, in which brain activity was to be measured in children as well as adults. The experiments sought to deploy cues across all three experiments that were less abstract, and hence more physically elaborate, than those used by Hill and Miller and by Lee *et al.*

Against this background, the three experiments presented in this chapter measured brain activity using electro-encephalography (EEG) in a two-talker listening task. Participants' task was to report key words spoken by a target talker in the presence of an opposite-gender competing talker that was presented simultaneously with, but from a contralateral spatial location as, the target talker. A visual cue was presented before the talkers spoke to inform participants, on each trial, about either the spatial location of the target talker (left/right of fixation) or their gender (male/female). The experiments aimed to identify robust attentional activity that did not reflect differences in physical aspects (e.g. luminance or complexity) of the visual stimuli used to cue attention. Therefore, a control condition was implemented to measure brain activity evoked by the visual cues in a condition in which they did not have implications for attention.

Both similarities and differences were expected to arise between the event-related potentials (ERPs) evoked by adults during attentional selection for location compared to gender. Similarities were expected to reflect domain-general processing of location and gender information, akin to the similarities in brain activity observed by Hill and Miller (2010) when listeners attended to talkers based on cues for location and F0. Differences in ERPs were expected to reflect cue-specific processing. Like Hill and Miller (2010) and Lee *et al.* (2013), the current experiments focussed on activity that arose in two phases of the task: (1) following the onset of the visual cue before the acoustic stimuli started (which is referred to as the "Preparatory Phase"), and (2) during the acoustic stimuli (referred to as the "Selective Phase").

Experiments 1 and 2 investigated EEG activity that was evoked by adults during two-talker listening, whereas Experiment 3 investigated whether domain-general and cue-specific activity was observed in children aged 7–13 years. A previous experiment by Dhamani, Leung, Carlile, and Sharma (2013) shows that, like adults, children benefit from advance cueing in noisy listening environments. Dhamani *et al.* asked children aged 10–15 years to identify a target syllable in a background of two-talker babble. On each trial, a cue was provided in advance of the target syllable to indicate the onset time of the target. Children were more accurate at identifying the target syllable when the cue validly predicted the onset time of the target, compared to when the target was presented earlier or later than expected. This result demonstrates that children aged 10–15 years are able to direct their attention to a talker based on a cue that indicates the onset time of the talker. Therefore, the ERPs evoked by children in Experiment 3 were expected to be similar to the ERPs evoked by adults in Experiments 1 and 2. Nevertheless, children typically identify speech in noise with lower accuracy than adults (Bonino, Leibold, & Buss, 2012; Fallon, Trehub, & Schneider, 2000; Papso & Blood, 1989). Therefore, it was expected that children aged 7–13 years would display weaker evidence than adults of significant EEG activity during the Preparatory and Selective Phases.

3.1. Experiment 1

3.1.1. Methods

3.1.1.1. Participants

Participants were 16 young adults (8 male), aged 18–24 years (mean [M] = 20.4, standard deviation [SD] = 1.5). They were self-declared right-handed native English speakers with no history of hearing problems. They had 5-frequency average pure-tone hearing levels of 20 dB HL or better, tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). The study was approved by the Research Ethics Committee of the Department of Psychology, University of York.

3.1.1.2. Apparatus

The experiment was conducted in a 5.3 m x 3.7 m single-walled test room (Industrial Acoustics Co., NY) located within a larger sound-treated room. Participants sat facing three loudspeakers (Plus XS.2, Canton) arranged in a circular arc at a height of 1 m at 0° azimuth (fixation) and at 30° to the left and right (Figure 3.1). A 15-inch visual display unit (VDU; NEC AccuSync 52VM) was positioned directly below the central loudspeaker.

3.1.1.3. Stimuli

Visual cues

Four visual cues, “left”, “right”, “male”, and “female”, were defined by white lines on a black background. Left and right cues were leftward- and rightward-pointing arrows, respectively; male and female cues were stick figures (Figure 3.2A–D). A composite visual stimulus consisted of the four cues overlaid (Figure 3.2E).

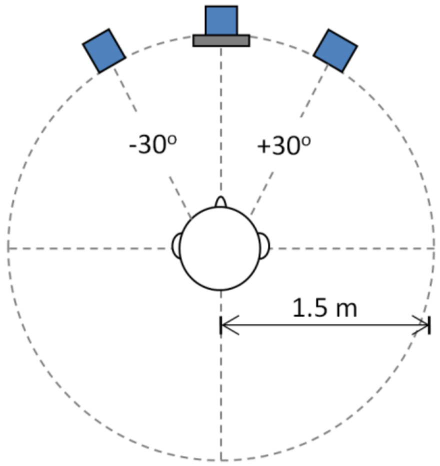


Figure 3.1. Layout of loudspeakers (blue squares) and visual display unit (grey rectangle) relative to a participant's head.

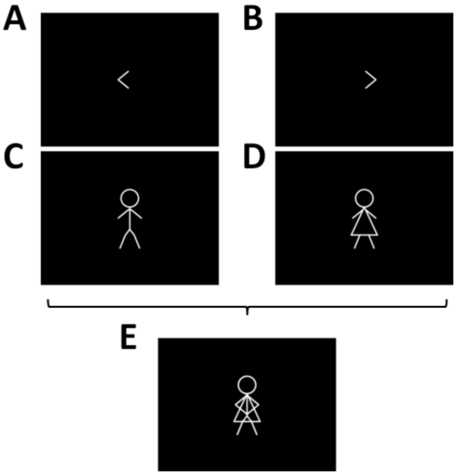


Figure 3.2. (**A**)-(**D**) Visual cues. (**E**) Visual composite stimulus, which is a combination of the four visual cues overlaid.

Acoustical test stimuli

Acoustical test stimuli were sentences from the Co-ordinate Response Measure corpus (CRM; Moore, 1981) spoken by native British-English talkers (Kitterick, Bailey, and Summerfield, 2010). CRM sentences have the form 'Ready <call sign>, go to <colour> <number> now'. In the sub-set used in the experiment, there were eight call-signs ('arrow', 'baron', 'charlie', 'eagle', 'hopper', 'laker', 'ringo', 'tiger'), four colours ('blue', 'red', 'green', 'white'), and four numbers ('1', '2', '3', '4'). An example is "Ready Charlie, go to Green Two now". Sentences spoken by three male talkers and three female talkers were selected from the corpus. The sentences had an average duration of 2.5 s. The levels of the digital recordings of the sentences were normalised to the same root mean square (RMS) power.

Acoustical control stimuli

Control stimuli were single-channel noise-vocoded representations of concurrent pairs of CRM sentences. Each control stimulus was created by summing a pair of sentences digitally with their onsets aligned, extracting the temporal envelope of the combination using the Hilbert Transform (Hilbert, 1912), and using the envelope to modulate the amplitude of random noise whose long-term spectrum matched the average spectrum of all of the pairs of sentences.

3.1.1.4. Procedures

Test Condition

At the start of each trial, a fixation cross was presented for 1000 ms. Next, the visual composite stimulus was presented. After 800 ms, parts of the composite stimulus began to fade, leaving only the visual cue for the trial. The fade lasted 200 ms and involved a decrease in luminance in order to minimise the onset response to the visual cue in the EEG recording. After the cue had been fully revealed for 1000 ms, two CRM sentences were presented concurrently. Two different sentences were presented from the two loudspeakers (left and right). The sentences started simultaneously, but contained different call signs and different colour-number combinations. The two talkers were selected quasi-randomly on each trial, with the restriction that one talker was male and the other was female. Over the course of the experiment, each of the six talkers was presented equally often from each location.

The visual cue directed attention to the target talker and varied quasi-randomly from trial to trial. The cue remained on the screen throughout the duration of the acoustic stimuli so that participants did not have to retain the visual cue in memory. After both sentences had ended, participants were instructed to report the colour-number combination in the target sentence by pressing a coloured digit on a touch screen directly in front of their chair. The inter-trial interval varied randomly from 1000 to 1500 ms to desynchronise anticipatory activity for the next trial. Each participant completed 384 trials (96 in each cueing condition), with a break every 48 trials.

The average presentation level of concurrent pairs of test sentences was set to 63 dB(A) SPL (range 61.6—66.2 dB) measured with a B&K (Brüel & Kjær, Nærum, Denmark) Sound Level Meter (Type 2260 Investigator) and 0.5-inch Free-field Microphone (Type 4189) placed in the centre of the arc at the height of the loudspeakers with the participant absent.

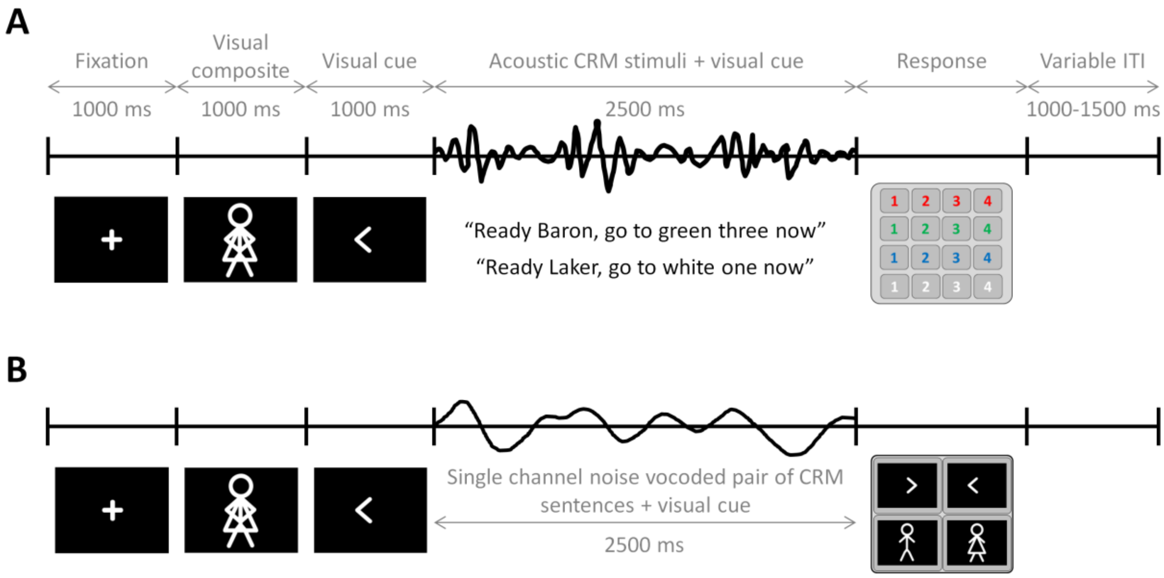


Figure 3.3. (**A**) Trial structure in the Test Condition, with an example trial below. (**B**) Trial structure in the Control Condition.

Control Condition

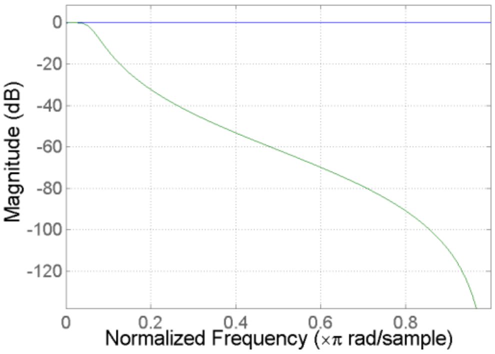
The trial structure of the Control Condition was identical to the Test Condition (Figure 3.3) with the exception that an acoustical control stimulus, presented from a single loudspeaker at 0° azimuth, replaced the pair of acoustical test stimuli. The task was to press the picture on the touch screen corresponding to the visual cue that was presented. Each participant completed 216 trials (54 in each visual stimulus condition), with a break every 36 trials. The presentation level of the control stimuli was set so that their average level matched the average level of the pairs of test stimuli. Participants undertook the Control Condition before the Test Condition; that is, before they had learnt the association between the visual cues and the acoustic stimuli.

The logic behind the design of the Control Condition was that the stimuli lacked the spectral detail and temporal fine structure required for the perception of pitch (Moore, 2008b). In addition, because the stimuli were presented from one loudspeaker, they did not provide the interaural differences in level and timing required for their constituent voices to be localised separately. In these ways, the acoustic cues required to segregate the sentences by gender and by location were neutralised, while the overall energy and gross fluctuations in amplitude of the test stimuli were preserved.

3.1.1.5. EEG recording and processing

Continuous EEG was recorded using the ANT WaveGuard-64 system (ANT, Netherlands; www.ant-neuro.com) with Ag/AgCl electrodes mounted on an elasticated cap (positions: Fp1, Fp2, AF3, AF4, AF7, AF8, F1, F2, F3, F4, F5, F6, F7, F8, FC1, FC2, FC3, FC4, FC5, FC6, FT7, FT8, C1, C2, C3, C4, C5, C6, T7, T8, CP1, CP2, CP3, CP4, CP5, CP6, TP7, TP8, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO7, PO8, O1, O2, M1, M2, Fpz, Fz, FCz, Cz, CPz, Pz, POz, Oz). An additional electrode (AFz) was used as a ground site. The horizontal electro-oculogram (EOG) was measured with a bipolar lead attached to the outer canthi of the left and right eyes and the vertical EOG was measured with a bipolar lead above and below the right eye. The EEG was amplified and digitised with an ANT High-Speed Amplifier (24 bit resolution) at a sampling rate of 1000 Hz per channel.

Figure 3.4. Graph showing properties of the EEG bandpass filter at different frequencies.



The continuous EEG recordings were exported to MATLAB 7 (The MathWorks, Inc., Natick, MA, USA) and analysed using the EEGLAB toolbox (Version 9; http://sccn.ucsd.edu/eeglab/). Before statistical analysis, the data were band-pass filtered between 0.25 and 30 Hz using a Butterworth filter (Figure 3.4). The amplitude at each electrode was referenced to the average amplitude of the electrode array. Epochs were created with 4700 ms duration, including a baseline interval of 200 ms at the end of the fixation-cross period. Epochs were rejected for further analysis if they contained high-amplitude artifacts (absolute amplitude in any channel greater than ± 200 μV) or if the behavioural response to the trial was incorrect. Independent component analysis (ICA) was used to correct for any remaining eye-blink artifacts, which were identified by a stereotyped scalp topography and a correlation with the vertical EOG that exceeded 0.6 for >80% of trials containing high-amplitude peaks.

3.1.1.6. Behavioural analyses

Trials were separated into Location (average left/right cues) and Gender (average male/female cues) groups, separately for the Test and Control Conditions. Responses were scored as correct if both the colour and number key words were reported correctly in the Test Condition, and if the visual cue was reported correctly in the Control Condition.

3.1.1.7. Analyses of ERPs

There were no expectations about where in the array of electrodes or when in time differences in ERPs between Test and Control conditions, or between Location and Gender trials, would occur during the Preparatory or Selective phases. Accordingly, in seeking significant differences, a Spatio-temporal Cluster-based Permutation Analysis was conducted (Maris & Oostenveld, 2007; see also Section 2.4.2).

The Spatio-temporal Cluster-based Permutation Analysis was used to make two types of comparison. Type-I analyses compared amplitudes in the Test and Control Conditions, separately for Location and Gender trials. Type-I clusters found in the Preparatory Phase could not arise from sensory or perceptual processes because the stimuli did not differ between the conditions in this phase. Rather, such differences were interpreted as arising from contrasting attentional activity between the Test and Control Conditions. Type-I clusters found in the Selective Phase, in contrast, could arise *either* from differences in attentional activity *or* from differences between the acoustical structure of the Test and Control stimuli.

Type-II analyses compared Location with Gender trials in the Test Condition. These analyses identified clusters where ERPs differed significantly depending on whether participants were receiving cues for, and directing attention towards, location or gender. Such differences could be evoked *either* by different attentional processes *or* by physical differences between the visual cues. Accordingly, we compared the average amplitude of Location and Gender trials—averaged over the space-by-time points in the cluster—between the Test and Control Conditions in a 2 x 2 ANOVA. The rationale was that differences in the visual cues between Location and Gender trials were also present in the Control Condition, but the attentional activity evoked by the cues should be present in the Test but not the Control Condition. A two-way interaction meant that the cluster could not be fully explained by the influence of physical differences in the visual cues between conditions. In order to determine whether such differences were sustained over the entire duration of a cluster or were restricted to particular moments, the difference of the difference in Location and Gender trials between the Test and Control Conditions was plotted, averaging only over the space-by-time points that fell in a 50-ms time window that was advanced in 10-ms steps over the duration of the cluster.

3.1.2. Results

3.1.2.1. Behavioural results

Conjoint accuracy in identifying the colour and number key words in the Test Condition was high and did not differ between Location (M = 95.3%, SD = 0.05) and Gender (M = 94.8%, SD = 0.05) trials, $t(15) = 1.2$, $p = 0.26$. There were also no significant differences in the accuracy with which the visual cue was identified in the Control Condition between Location (M = 99.4%, SD = 0.01) and Gender (M = 99.1%, SD = 0.02) trials, $t(15) = 0.7$, $p = 0.51$.

3.1.2.2. Event-related potentials

Type-I analyses: Differences between Test and Control Conditions

Location trials

Figure 3.5 illustrates the results of the Type-I analyses on trials in which a Location cue (left/right) was presented. The latencies of significant clusters are plotted relative to the onset of the talkers (i.e. relative to the start of the Selective Phase). The descriptions in the following paragraphs describe the latencies of significant clusters relative to the start of the phase in which the cluster occurred (i.e. the latencies of clusters that occurred during the Preparatory Phase are reported relative to the start of the Preparatory Phase)[7].

During the 1000-ms Preparatory Phase, one significant cluster of activity (Cluster 1) was identified (Figure 3.5A). It involved 25 central electrodes and spanned the time interval

---

[7] The same logic applies to the results described throughout this chapter and for the remainder of the thesis.

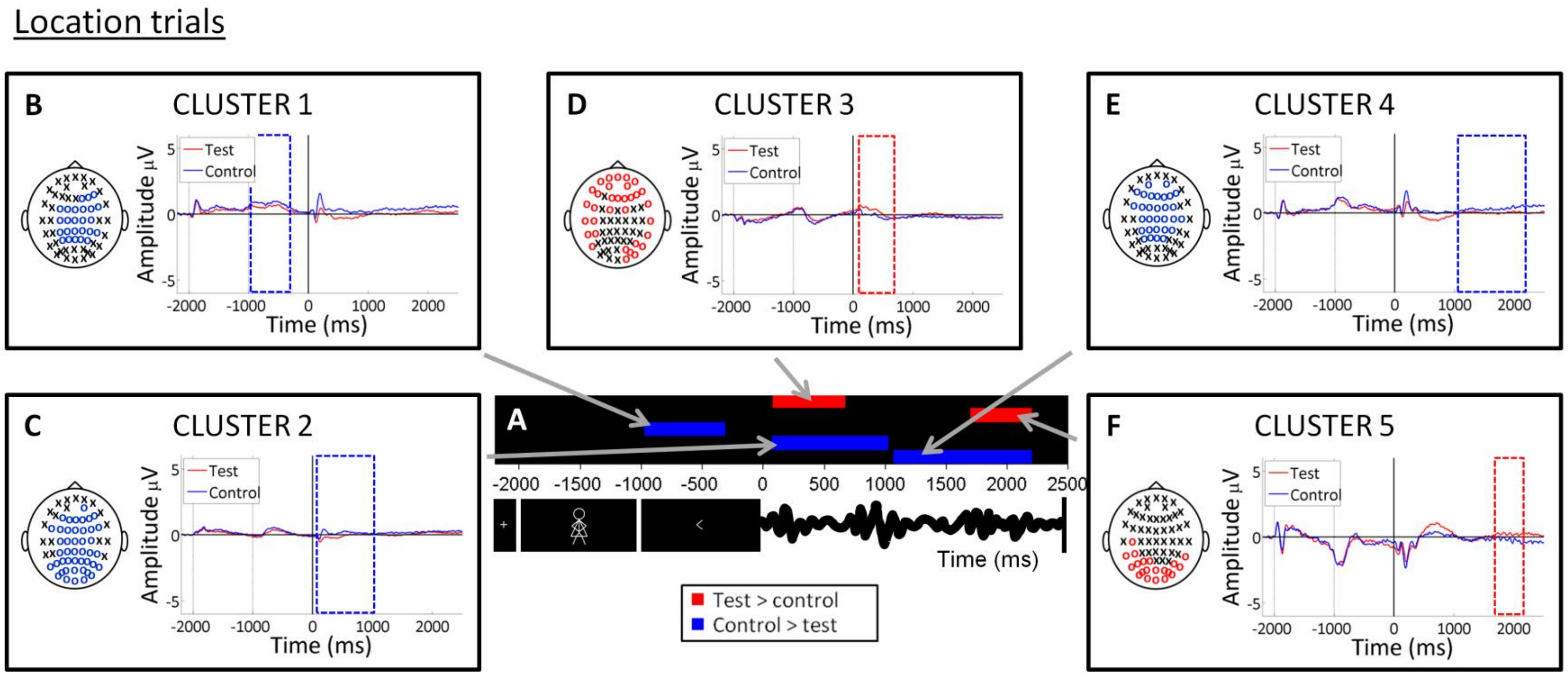


Figure 3.5. (Continued on next page). Experiment 1: Results from Type-I Spatio-temporal Cluster-based Permutation Analyses for Location (**A** to **F**) and Gender (**G** to **J**) trials. (**A** and **G**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Further information about each cluster is displayed in (**B** to **F** and **H** to **J**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

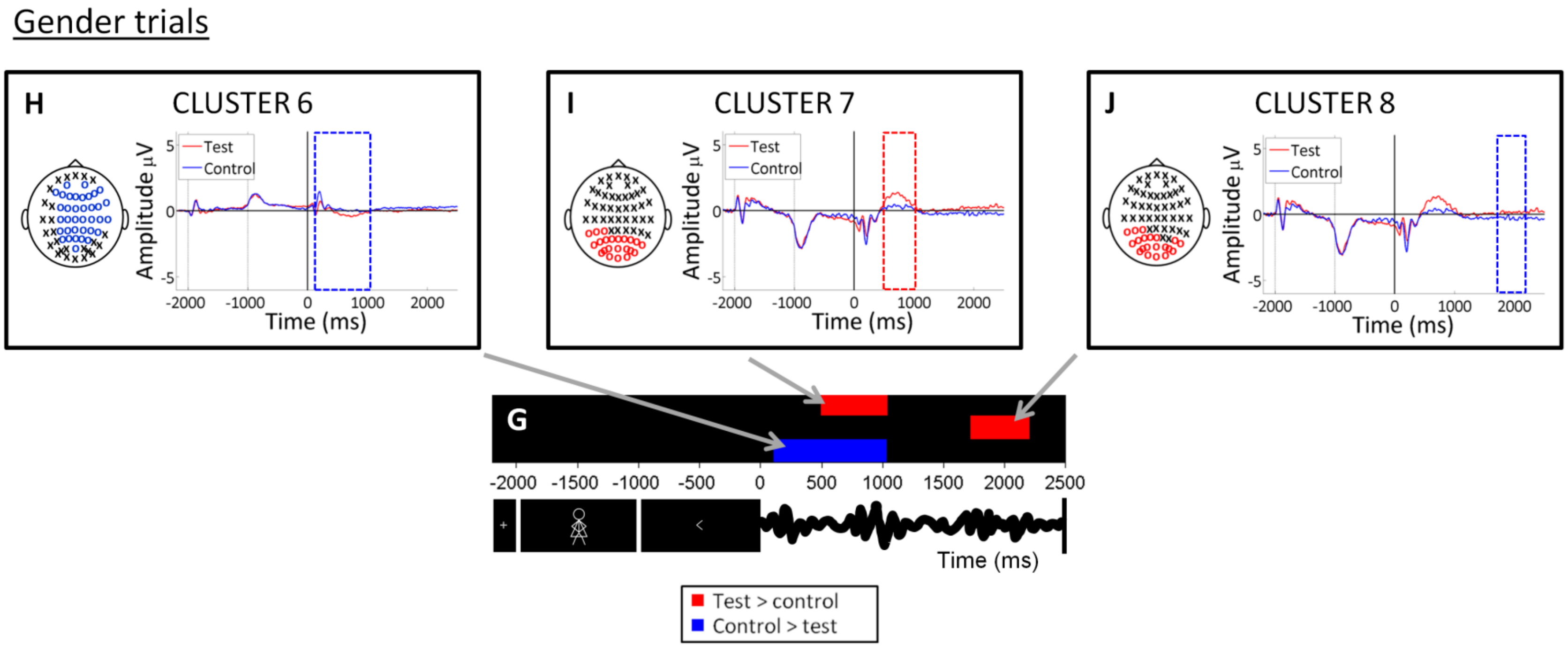


Figure 3.5. (Continued from the previous page)

Table 3.1. (Continued on next page). Summary of results for the Gender and Location Condition comparisons (Type-I analysis) between the Test and Control Conditions across Experiments 1–3. The rows headed 'Cluster p-value' show the results of the Spatio-temporal Cluster-based Permutation Analysis.

Phase		Experiment 1 Location	Experiment 2 Location	Experiment 3 Location	Experiment 1 Gender	Experiment 2 Gender	Experiment 3 Gender
Preparatory	**Cluster Number**	**1**	**14**	**25**	-	**17**	-
	Cluster *p*-value	0.001	0.014	0.039	-	0.03	-
	Polarity	Control > Test	Control > Test	Control > Test	-	Control > Test	-
	Electrode Locations	Central	Central	Posterior	-	Central	-
	Onset of cluster (ms)	27	43	210	-	84	-
	Duration of cluster (ms)	664	638	245	-	300	-
Selective	**Cluster Number**	**2**	**15**	**26**	**6**	**18**	**29**
	Cluster *p*-value	0.002	0.001	< 0.001	0.002	< 0.001	0.010
	Polarity	Control > Test	Control > Test	Control > Test	Control > Test	Control > Test	Control > Test
	Electrode Locations	Posterior + Central	Central	Posterior + Central	Central + Anterior	Central	Central
	Onset of cluster (ms)	69	286	71	108	112	573
	Duration of cluster (ms)	960	1176	1432	922	1107	494
Selective *(continued on next page)*	**Cluster Number**	**3**	**16**	**27**	**7**	**19**	**28**
	Cluster *p*-value	0.041	0.014	0.034	0.010	0.007	0.021
	Polarity	Test > Control	Test > Control	Test > Control	Test > Control	Test > Control	Test > Control

Table 3.1. (Continued from previous page)

Phase		**Experiment 1 Location**	**Experiment 2 Location**	**Experiment 3 Location**	**Experiment 1 Gender**	**Experiment 2 Gender**	**Experiment 3 Gender**
Selective *(continued from previous page)*	**Electrode Locations**	Non-Central	Posterior	Non-Central	Posterior	Posterior	Posterior
	Onset of cluster (ms)	81	298	77	495	502	483
	Duration of cluster (ms)	590	704	864	543	529	556
Selective	**Cluster Number**	**4**	-	-	-	**20**	-
	Cluster *p*-value	0.002	-	-	-	0.003	-
	Polarity	Control > Test	-	-	-	Control > Test	-
	Electrode Locations	Central	-	-	-	Central	-
	Onset of cluster (ms)	1072	-	-	-	1261	-
	Duration of cluster (ms)	1128	-	-	-	939	-
Selective	**Cluster Number**	**5**	-	-	**8**	**21**	-
	Cluster *p*-value	0.027	-	-	0.033	0.047	-
	Polarity	Test > Control	-	-	Test > Control	Test > Control	-
	Electrode Locations	Posterior	-	-	Posterior	Posterior	-
	Onset of cluster (ms)	1696	-	-	1717	1844	-
	Duration of cluster (ms)	504	-	-	483	356	-

from 27 to 691 ms, relative to the start of the phase. It showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 16699, p = 0.001] (Figure 3.5B). The existence of Cluster 1 demonstrates that differences in brain activity arise between a condition in which a visual cue has no implications for auditory attention and a condition in which the same visual cue directs listeners to prepare to select an upcoming talker on the basis of their location. The differences in brain activity arose 227 ms after the visual cue began to appear and 27 ms after the visual cue was fully revealed. The polarity, location, onset time, and duration of Cluster 1 are tabulated in the third column of Table 3.1.

During the Selective Phase, four significant clusters of activity were identified (Figure 3.5A). Cluster 2 (Figure 3.5C) spanned the interval from 69 to 1029 ms, relative to the start of the phase. It involved 44 central and posterior electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 48457, *p* = 0.002]. Cluster 3 (81 to 671 ms; Figure 3.5D) was complementary to Cluster 2 since the time points of these clusters overlapped. Cluster 3 involved 33 non-central electrodes and showed significantly more positive amplitude during the Test Condition than the Control Condition [cluster statistic = 13476, *p* = 0.041]. Cluster 4 (1072 to 2200 ms; Figure 3.5E) started shortly after Cluster 2 had finished. It involved 30 central electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 44288, *p* = 0.002]. Cluster 5 (1696 to 2200 ms; Figure 3.5F) started towards the end of the Selective Phase. It involved 20 posterior electrodes and showed significantly more positive amplitude during the Test Condition than the Control Condition [cluster statistic = 16619, *p* = 0.027].

Gender trials

The second of the Type-I analyses compared ERPs between the Test and Control Conditions on trials in which a Gender cue (male/female) was presented. Panels G–J of Figure 3.5 illustrate these results. No significant clusters were identified during the Preparatory Phase. During the Selective Phase, three significant clusters were identified (Figure 3.5G). Cluster 6 (108 to 1030 ms; Figure 3.5H) involved 36 central and anterior electrodes. It showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 35606, *p* = 0.002]. Cluster 7 (495 to 1038 ms; Figure 3.5I) was complementary to Cluster 6. It involved 22 mainly posterior electrodes which showed significantly more positive amplitude during the Test Condition than the Control Condition [cluster statistic = 24580, *p* = 0.010]. Cluster 8 (1717 to 2200 ms; Figure 3.5J) occurred later during the Selective Phase. It involved 20 mainly posterior electrodes and displayed significantly more positive amplitude during the Test Condition than the Control Condition

Table 3.2. (Continued on next page). Summary of results for the Test Condition comparison between Location and Gender trials (Type-II analysis) across Experiments1–3. A tick in the row headed 'Significant in Control Condition?' indicates that the difference in the amplitude of ERPs between Location and Gender trials was significant in the Control Condition across the spatio-temporal points of the cluster (p-values displayed underneath). A tick in the row headed 'Significant Test/Control Interaction?' indicates that an ANOVA with the factors cue type (Location/Gender) and condition (Test/Control) revealed a significant two-way interaction (p-values displayed underneath).

Phase	Properties	Experiment 1	Experiment 2	Experiment 3
Preparatory	**Cluster Number**	**9**	**22**	**30**
	Cluster p-value	< 0.001	0.004	0.014
	Polarity	Loc > Gen	Loc > Gen	Loc > Gen
	Electrode Locations	Posterior	Posterior	Posterior
	Onset of cluster (ms)	29	53	72
	Duration of cluster (ms)	599	342	372
	Significant in Control Condition?	✓ $p = 0.011$	✓ $p = 0.017$	✓ $p < 0.001$
	Significant Test/Control Interaction?	✗ $p = 0.82$	✗ $p = 0.85$	✓ $p = 0.003$
Preparatory	**Cluster Number**	**10**	**23**	-
	Cluster p-value	0.003	0.005	-
	Polarity	Gen > Loc	Gen > Loc	-
	Electrode Locations	Anterior + Central	Anterior + Central	-
	Onset of cluster (ms)	40	103	-
	Duration of cluster (ms)	389	288	-
	Significant in Control Condition?	✓ $p = 0.002$	✓ $p = 0.014$	-
	Significant Test/Control Interaction?	✗ $p = 0.80$	✗ $p = 0.80$	-
Preparatory	**Cluster Number**	**11**	-	-
	Cluster p-value	0.010	-	-
	Polarity	Gen > Loc	-	-
	Electrode Locations	Central	-	-
	Onset of cluster (ms)	484	-	-
	Duration of cluster (ms)	464	-	-
	Significant in Control Condition?	✗ $p = 0.15$	-	-
	Significant Test/Control Interaction?	✗ $p = 0.90$	-	-

Table 3.2. (Continued from the previous page)

	Cluster Number	**12**	**24**	**31**
Selective	**Cluster *p*-value**	< 0.001	0.018	0.022
	Polarity	Gen > Loc	Gen > Loc	Gen > Loc
	Electrode Locations	Posterior + Central	Central	Central
	Onset of cluster (ms)	371	807	1069
	Duration of cluster (ms)	835	396	455
	Significant in Control Condition?	× p = 0.56	× p = 0.31	× p = 0.07
	Significant Test/Control Interaction?	× p = 0.08	✓ p = 0.044	✓ p = 0.001

	Cluster Number	**13**	-	-
Selective	**Cluster *p*-value**	0.049	-	-
	Polarity	Loc > Gen	-	-
	Electrode Locations	Anterior	-	-
	Onset of cluster (ms)	590	-	-
	Duration of cluster (ms)	279	-	-
	Significant in Control Condition?	× p = 0.26	-	-
	Significant Test/Control Interaction?	✓ p = 0. 005	-	-

[cluster statistic = 14722, p = 0.033]. Many of the electrodes in Cluster 8 overlapped with the electrodes that contributed to Cluster 7.

Type-II analyses: Differences between Location and Gender trials

Differences during the Preparatory Phase

During the Preparatory Phase, three clusters of electrodes were identified that differed significantly in the Test Condition between Location and Gender trials (Figure 3.6A). Cluster 9 (29 ms to 628 ms; Figure 3.6B) involved 28 mainly posterior electrodes and showed significantly more positive amplitude during Location trials than Gender trials [cluster statistic = 21609, p < 0.001]. These values are listed in the third column of Table 3.2. For this cluster, the interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [F(1,15) = 0.05, p = 0.82; Figure 3.7A] and the difference between Location and Gender trials was also present in the Control Condition, p = 0.011. When the difference of the differences in Location and Gender trials between the Test and Control conditions was examined in 50-ms sliding windows, the uncorrected p-value was less than 0.05 in only seven of the 60 50-ms time windows in the cluster (Figure 3.7F). The finding that ERPs in this cluster did not differ between the Test and Control Conditions means that it is not possible to

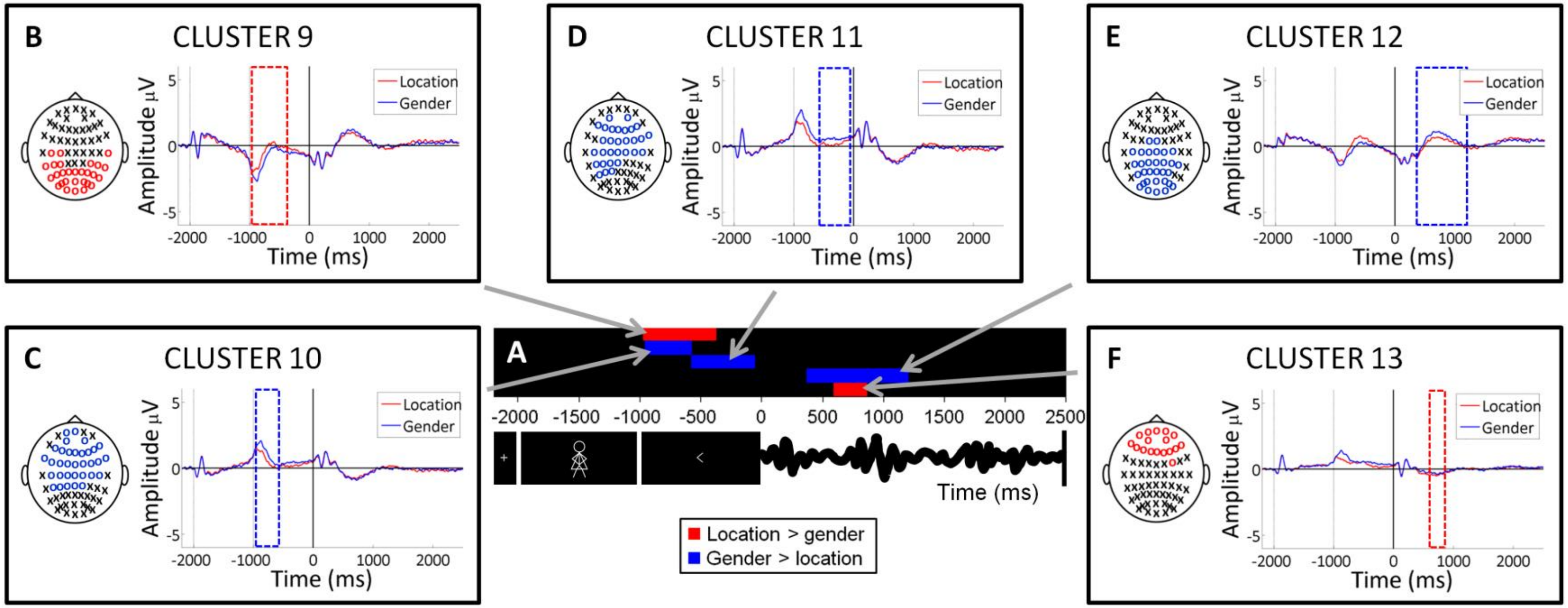


Figure 3.6. Experiment 1: Results of the Type-II Spatio-temporal Cluster-based Permutation Analyses. This analysis contrasted Location and Gender trials in the Test Condition. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive on Location trials than Gender trials. For clusters plotted as blue rectangles, the average amplitude was more negative on Location trials than Gender trials. Further information about each cluster is displayed in (**B**)-(**F**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

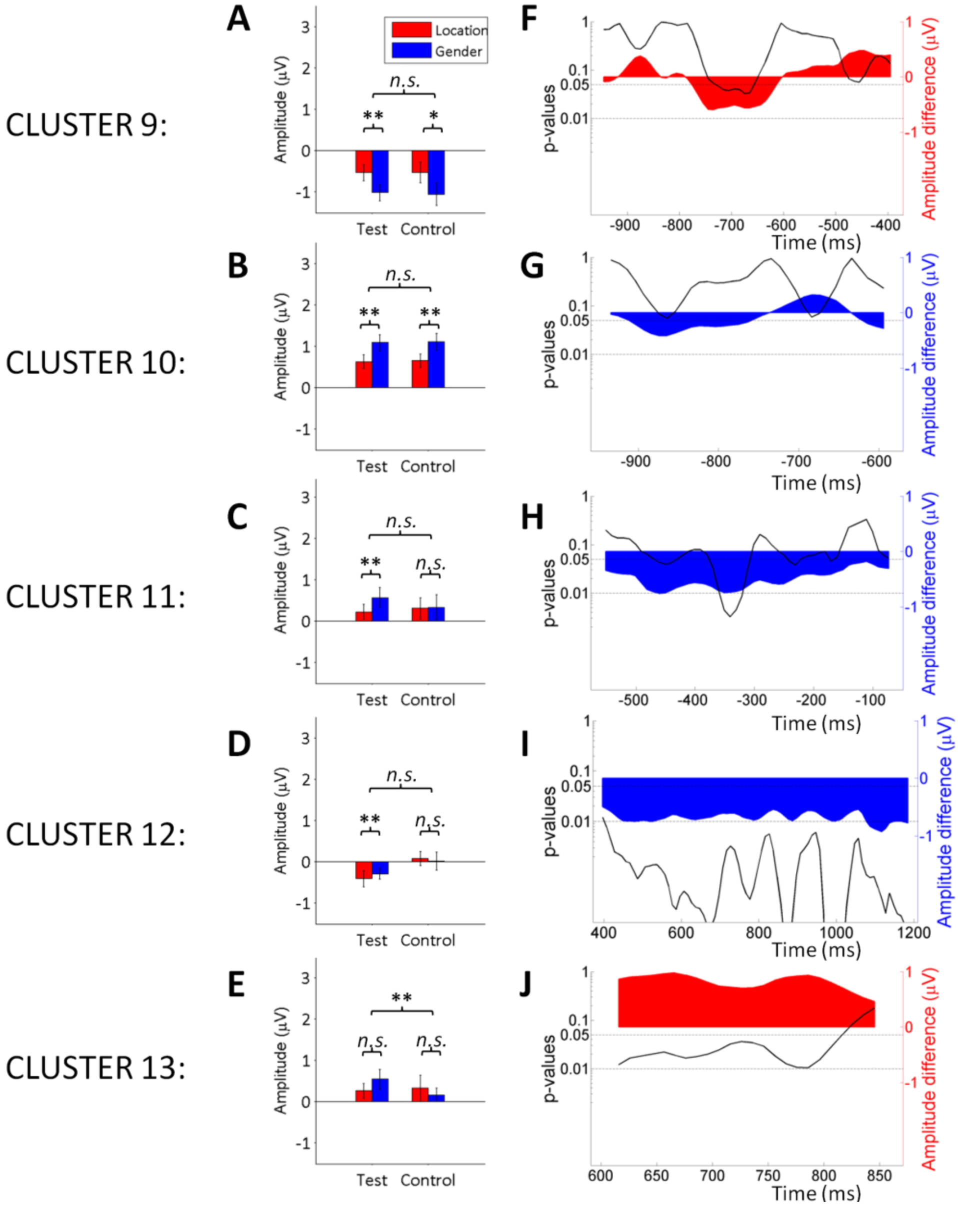


Figure 3.7. Experiment 1: Comparison of differences in the amplitude of ERPs between Location and Gender trials in the Test and Control Conditions for each significant Type-II cluster in Experiment 1. Graphs (**A**)-(**E**) plot the mean amplitude for Location and Gender trials in the Test and Control Conditions, averaged across participants and space-time points. Error bars show 95% within-subjects confidence intervals. Narrow brackets display the significance level of the comparison between Location and Gender trials in the Test and Control Conditions. Wider brackets display the significance level of the two-way interaction (* $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). Graphs (**F**)-(**J**) display the difference of the differences in Gender and Location trials between the Test and Control conditions in 50-ms time windows repeated every 10 ms within the cluster (right axis) and the uncorrected *p*-values resulting from a paired-samples *t*-test comparing the differences (left axis). The mid-point of each time window relative to the onset of acoustic stimuli is displayed on the x-axis.

rule out the explanation that the cluster arose from differences in the visual cues, rather than from differences in attentional processes triggered by the cues.

Cluster 10 (Figure 3.7C) was complementary to Cluster 9 and was also likely to arise from differences in the visual cues. Cluster 10 (40 to 429 ms) started slightly after Cluster 9, but involved a largely complementary group of electrodes that displayed amplitudes of opposite polarity. It involved 33 central and anterior electrode locations and showed significantly more negative amplitude in Location trials than Gender trials (cluster statistic = 188274; p = 0.003). For Cluster 10, like Cluster 9, the interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [$F(1,15)$ = 0.01, p = 0.80; Figure 3.7B] and the difference between Location and Gender trials was also present in the Control Condition, p = 0.002. In addition, the uncorrected p-value did not fall below 0.05 during any 50-ms segment over the duration of the cluster (Figure 3.7G).

Cluster 11 (484 to 948 ms; Figure 3.6D) arose later during the Preparatory Phase. It showed significantly more negative amplitude in Location trials than Gender trials [cluster statistic = 120364, p = 0.010] and some of the electrodes overlapped with those identified in Cluster 2. The interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [$F(1,15)$ = 2.32, p = 0.15; Figure 3.7C], although the difference between Location and Gender trials was not significant in the Control Condition, p = 0.90. Figure 3.7H shows that the difference between the Test and Control Conditions reached the $p < 0.001$ (uncorrected) criterion in 50-ms segments centred between 650 and 680 ms. The finding that ERPs did not differ between Location and Gender trials in the Control Condition implies that activity within this cluster might reflect differences in the attentional processes triggered by the cues. However, the finding of no significant interaction means that it was not possible to fully rule out the explanation that the cluster arose from differences in the visual cues.

Differences during the Selective Phase

During the Selective Phase, two clusters of activity were identified that differed significantly between Location and Gender trials (Figure 3.6A). Cluster 12 (Figure 3.6E) lasted from 371 to 1206 ms after the start of the acoustic stimuli. It involved 31 central and posterior electrode locations and displayed significantly more negative amplitude in Location trials than Gender trials [cluster statistic = 25506, p < 0.001]. The interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [$F(1,15)$ = 3.46, p = 0.08; Figure 3.7D], although the difference between Location and Gender trials was not significant in the Control Condition, p = 0.56. This pattern of amplitudes in the Control Condition is similar to Cluster 11. However, for Cluster 12, the p-value for the difference between Test and Control Conditions was below 0.05 (uncorrected) in 50-ms windows throughout the cluster (Figure 3.7I).

Cluster 13 (Figure 3.6F) started after Cluster 12 but overlapped it in time. Cluster 13 lasted from 590 to 869 ms after the onset of the acoustic stimuli. It involved 17 anterior electrode locations and showed significantly more positive amplitude in Location than Gender trials [cluster statistic = 6501, $p = 0.049$]. There was a significant interaction between cue type (Location/Gender) and condition (Test/Control) [$F(1,15) = 11.07$, $p = 0.005$; Figure 3.7E] and the difference between Location and Gender trials was not significant in the Control Condition, $p = 0.26$. The finding of a significant interaction demonstrates that Cluster 13 arose from differences in the processes for attending selectively to a talker between Location and Gender trials. In addition, the *p*-value for the difference between the Test and Control Conditions was below 0.05 (uncorrected) in 50-ms windows over most of the duration of the cluster (Figure 3.7J).

3.1.3. Discussion

During the Preparatory Phase, Type-I analyses demonstrated that significantly different ERPs occurred in the Test Condition compared with the Control Condition, but only on Location trials (Figures 3.5A and B) and not on Gender trials (Figure 3.5G). During the Preparatory Phase, no acoustical stimuli had been presented and the visual stimuli did not differ between the Test and Control Conditions. The result indicates, therefore, that listeners evoke different brain activity when a visual cue indicates the location of the target talker than when the same cue has no implications for auditory attention. The result is compatible with the interpretation that the visual cue can trigger preparatory attentional activity. Moreover, it does so with a short latency (< 30 ms) after the full reveal of the visual cue.

The Type-II analysis in the Test Condition revealed significant differences between Location and Gender trials in the Preparatory Phase, with a similar latency as Type-I differences during the Location Condition (Figures 3.6B–C). However, a difference between Location and Gender trials also occurred in the Control Condition at the same electrodes and time points (Figures 3.7A–B). Thus, it is not possible to rule out the explanation that these early clusters were evoked largely by physical differences between the visual cues for location compared with gender, rather than by differences in preparatory attentional processes triggered by the different cue types. The physical differences may have involved luminance and structural complexity. A further component of the difference in ERPs may have arisen from differences in the cognitive processes evoked by the representation of an inanimate object (a chevron) compared with a human being (Caramazza & Shelton, 1998; Downing, Chan, Peelen, Dodds, & Kanwisher, 2006).

The behavioural results demonstrate that participants could correctly identify words spoken by the target talker in both Location and Gender trials, even though there was no evidence of preparatory EEG activity in Gender trials. This outcome could have arisen from a feature our design. Whereas there were only two possible locations, there were three possible male and three possible female talkers. As a result, there was more variation in the evidence of gender (e.g. in average values of the F0 and formant frequencies) than in the evidence of location. Thus, the cues for location were more specific than the cues for gender. Even though the difference in specificity was not reflected in differences in behavioural accuracy, it might have influenced the patterns of brain activity that were observed during the Preparatory Phase. Experiment 2 tested two hypotheses: first, that gender cues evoke preparatory brain activity when variation in the evidence of gender is minimised, and second, that differential activity emerges between Location and Gender trials when both types of cue are similarly specific. An additional aim was to determine whether the overall pattern of results of Experiment 1 could be replicated with a different set of participants.

3.2. Experiment 2

To avoid differences in the specificity of the visual cues for attributes of the target talker between Location and Gender trials, the same male and female talker were presented for the entire experiment, rather than employing three instances of each gender as in Experiment 1. Also, participants were familiarised with the locations and genders before the Test Condition was administered.

3.2.1. Methods

3.2.1.1. Participants

Participants were 16 young adults (8 male), aged 18–27 (M = 21.3, SD = 2.1), none of whom had taken part in Experiment 1. All participants were self-declared right-handed native English speakers with no history of hearing problems. Participants all had 5-frequency average pure-tone hearing levels of 20 dB HL or better, tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). The study was approved by the Research Ethics Committee of the Department of Psychology, University of York.

3.2.1.2. Stimuli and procedure

Stimuli and procedures were the same as those in Experiment 1 except that only one of the male and one of the female talkers were used. After participants had completed the Control Condition, but before they undertook the Test Condition, a set of trials aimed to familiarise participants with the two locations and the two talkers. Familiarisation involved 52 trials in which only one or other of the two talkers, but not both, was presented during the Selective Phase. The trial structure was identical to the Test Condition except that there was no competing talker and EEG was not recorded.

3.2.1.3. EEG recording, processing, and analyses

The EEG recording, processing, and analysis procedures were identical to those used in Experiment 1.

3.2.2. Results

3.2.2.1. Behavioural results

Conjoint accuracy in identifying the colour and number key words in the Test Condition was high and did not differ between Location (M = 96.5%, SD = 0.02) and Gender (M = 95.9%, SD = 0.02) trials, $t(15) = 1.0$, $p = 0.34$. There were also no significant differences in the accuracy with which the visual cue was identified in the Control Condition between Location (M = 99.6%, SD = 0.01) and Gender (M = 99.6%, SD = 0.01) trials, $t(15) = 0.3$, $p = 0.79$.

3.2.2.2. Event-related potentials

Type-I analyses: Differences between Test and Control Conditions

Location trials

Figure 3.8 illustrates the results of the Type-I analyses. Panels A–D report analyses that compared ERPs between the Test and Control Conditions on trials in which a Location cue was presented. One significant cluster of activity was identified during the Preparatory Phase (Figure 3.8B) and two significant clusters were identified during the Selective Phase (Figure 3.8C–D). The polarity, location, onset time, and duration of these clusters are listed in Table 3.1.

Gender trials

Panels E–J of Figure 3.8 illustrate the results of the Type-I analysis that compared ERPs between the Test and Control Conditions on trials in which a Gender cue was presented. One significant cluster was identified during the Preparatory Phase (Figure 3.8F) and four significant clusters were identified during the Selective Phase (Figure 3.8G–J). The polarity, location, onset time, and duration of these clusters are listed in Table 3.1.

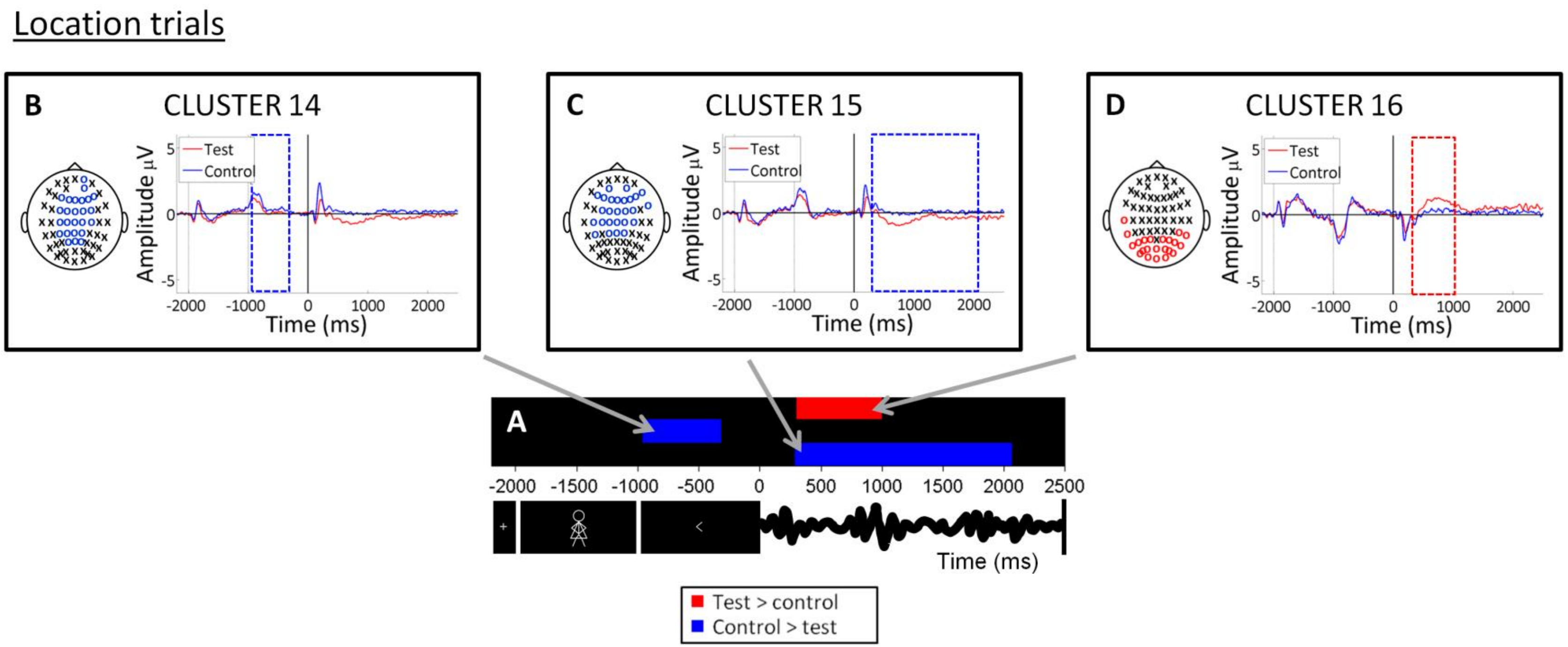


Figure 3.8. (Continued on next page). Experiment 2: Results from Type-I Spatio-temporal Cluster-based Permutation Analyses for the Location Condition (**A** to **F**) and the Gender Condition (**E** to **J**). (**A** and **E**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Further information about each cluster is displayed in (**B** to **D** and **F** to **J**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

Gender trials

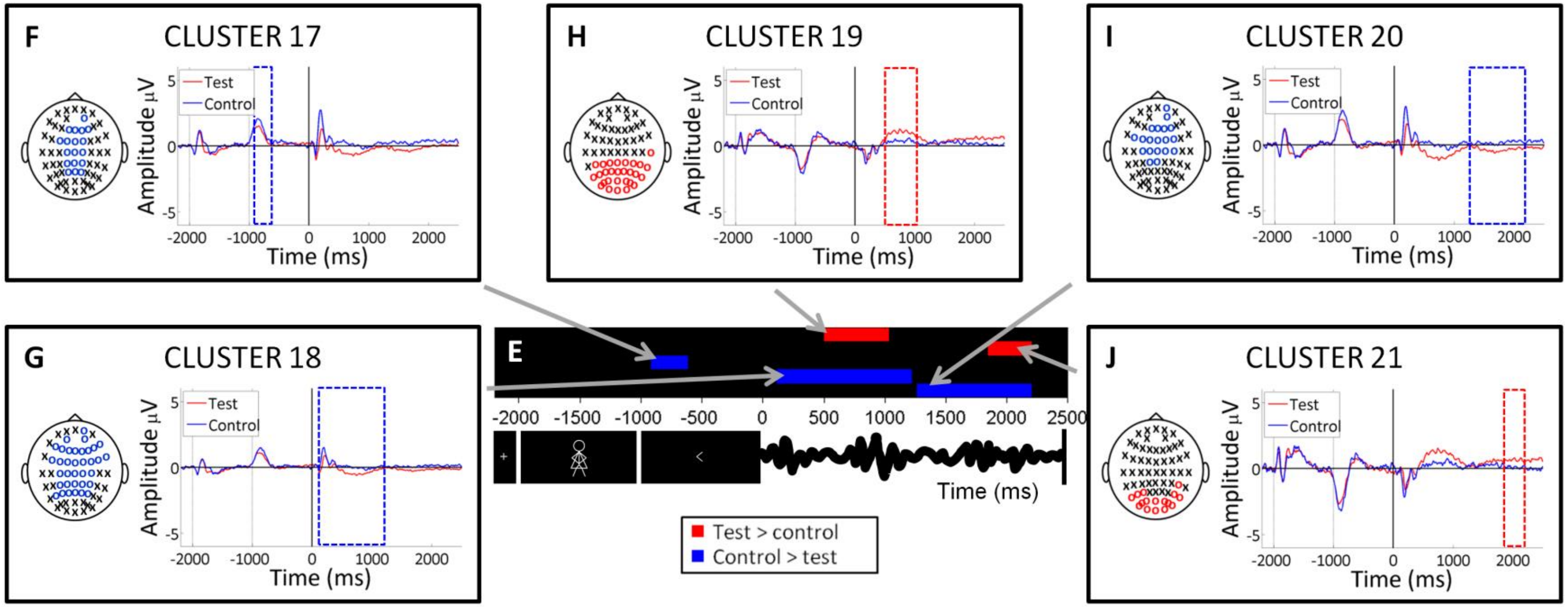


Figure 3.8. (Continued from the previous page)

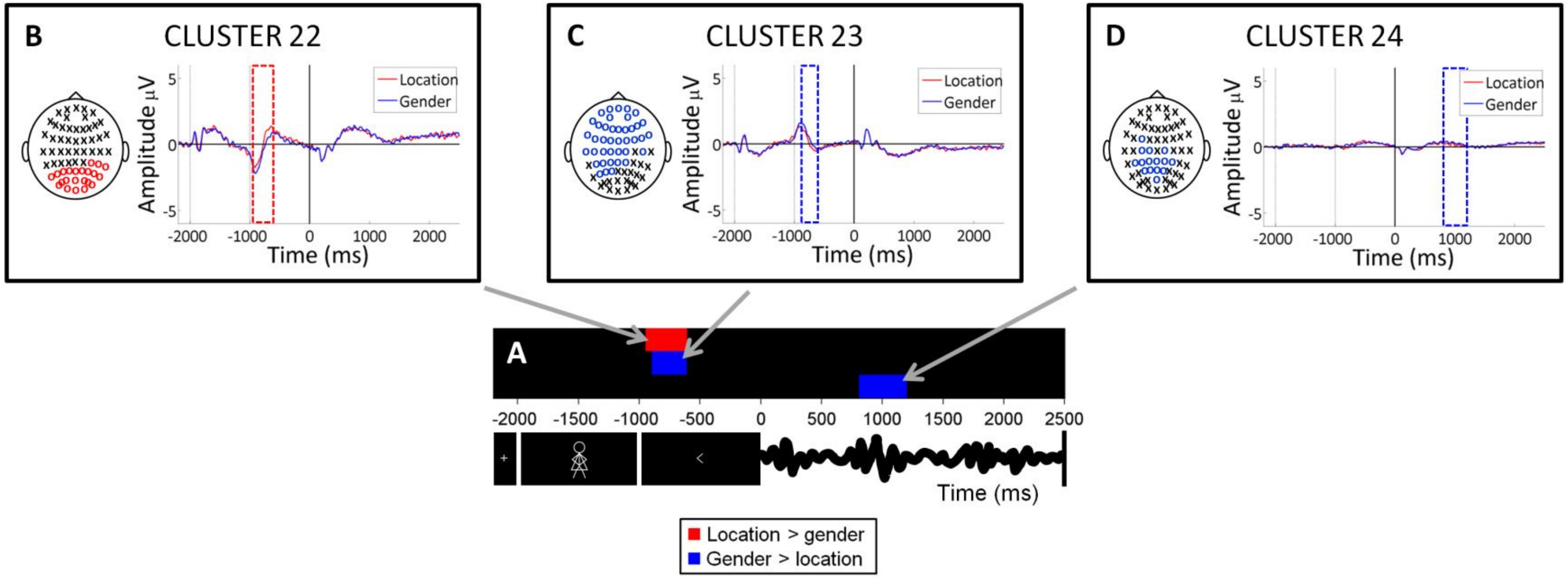


Figure 3.9. Experiment 2: Results from the Type-II Spatio-temporal Cluster-based Permutation Analysis. This analysis contrasted Location and Gender trials in the Test Condition. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive on Location trials than Gender trials. For clusters plotted as blue rectangles, the average amplitude was more negative on Location trials than Gender trials. Further information about each cluster is displayed in (**B**)-(**D**). For each cluster, the topographical map shows the electrodes that contribute to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

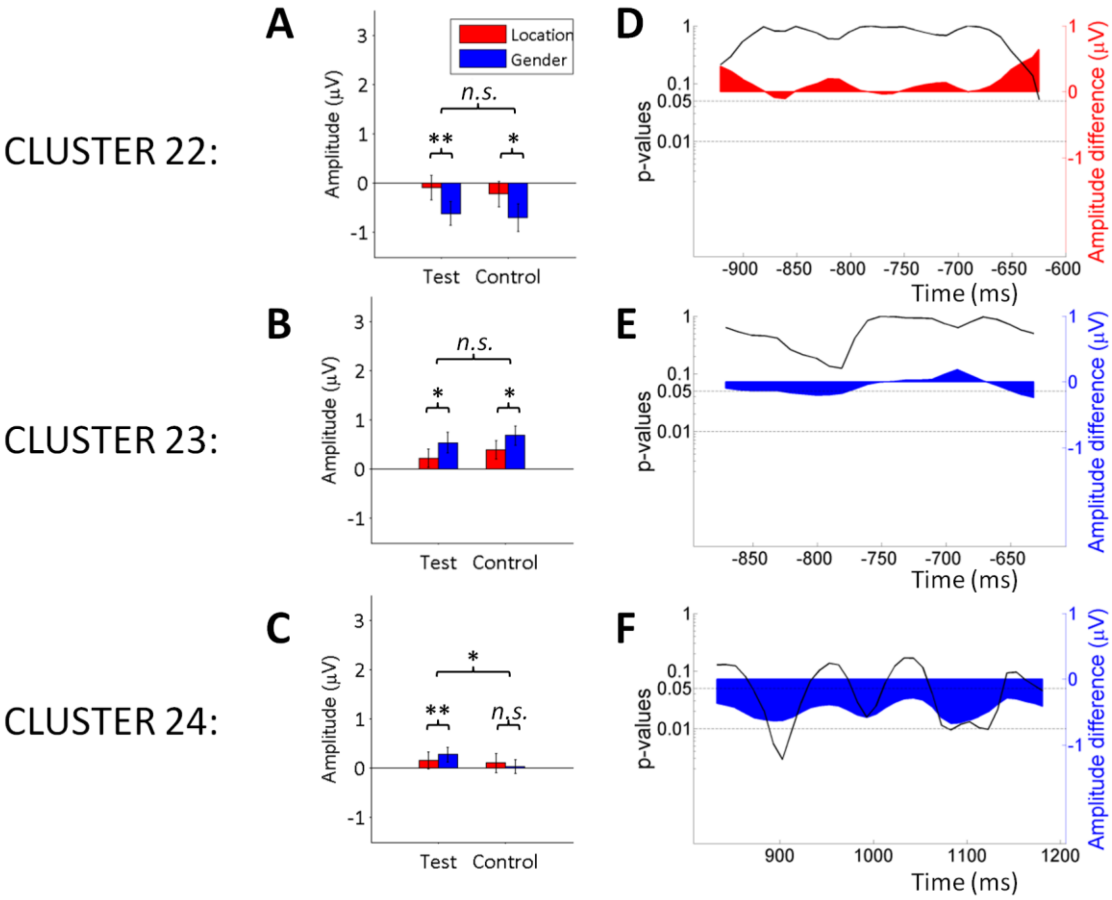


Figure 3.10. Experiment 2: Comparison of differences in the amplitude of ERPs between Location and Gender trials in the Test and Control Conditions for each significant Type-II cluster in Experiment 1. Graphs (**A**)-(**C**) plot the mean amplitude for Location and Gender trials in the Test and Control Conditions, averaged across participants and space-time points. Error bars show 95% within-subjects confidence intervals. Narrow brackets display the significance level of the comparison between Location and Gender trials in the Test and Control Conditions. Wider brackets display the significance level of the two-way interaction (* $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). Graphs (**D**)-(**F**) display the difference of the differences in Gender and Location trials between the Test and Control conditions in 50-ms time windows repeated every 10 ms within the cluster (right axis) and the uncorrected p-values resulting from a paired-samples t-test comparing the differences (left axis). The mid-point of each time window relative to the onset of acoustic stimuli is displayed on the x-axis.

Type-II analyses: Differences between Location and Gender Conditions

Figure 3.9 illustrates the results of Type-II analyses that compared ERPs between Location and Gender trials in the Test Condition. The analysis identified two significant clusters during the Preparatory Phase (Figure 3.9B–C) and one significant cluster during the Selective Phase (Figure 3.9D). The polarity, location, onset time, and duration of these clusters are listed in Table 3.2.

The clusters identified during the Preparatory Phase (Clusters 22 and 23) showed the same patterns of activity in the Control Condition ($p \leq 0.017$; Figure 3.10). For these clusters, the interaction between cue type (Location/Gender) and condition (Test/Control) was not significant. Therefore, it is not possible to rule out the explanation that Type-II clusters during the Preparatory Phase arose from differences in the visual cues, rather than from differences in attentional proceses triggered by the cues. When the difference between Test and Control Conditions was examined in 50-ms sliding windows, the p-values did not fall below the 0.05 (uncorrected) level at any time point during any of the clusters.

The cluster identified during the Selective Phase (Cluster 24) did not show the same pattern in the Control Condition ($p = 0.31$; Figure 3.10C). The interaction between cue type (Location/Gender) and condition (Test/Control) was significant [$F(1,15) = 4.82$, $p = 0.044$]. This finding demonstrates that the cluster during the Selective Phase arose from differences in the processes for attending selectively to a talker between Location and Gender trials. Figure 3.10F shows that the *p*-values were less than the 0.05 (uncorrected) level at several of the 50-ms segments over the duration of the cluster—first around 900 ms, then around 1000 ms, and finally between 1060 and 1140 ms.

3.2.3. Discussion

Experiment 2 partially replicated the results of Experiment 1. Both experiments provide evidence for activity during the Preparatory Phase of Location trials that began earlier than 50 ms after the visual cue was fully revealed, lasted longer than 600 ms, and was characterised by more negative amplitudes for the Test than Control Condition at central electrodes (Figures 3.5B and 3.8B). Additionally, both experiments revealed Type-II differences between Location and Gender trials during the Preparatory Phase that were present both in the Test and in the Control Conditions. The findings during the Selective Phases are also similar. Type-I differences in Location and Gender trials occurred throughout the Selective Phase, characterised by more negative amplitudes during the Test than Control Condition at central electrodes and more positive amplitudes at non-central (typically posterior) electrodes. Type-II results revealed more negative amplitudes on Location than Gender trials at central electrodes during the Selective Phase, which were not present in the Control Condition.

A difference between the experiments is the finding of significant activity during the Preparatory Phase of Gender trials in Experiment 2, which was not present in Experiment 1. One interpretation is that presenting the same male and female talker throughout the experiment enabled participants to engage preparatory attention in response to gender cues in ways that were not possible when the set of talkers was larger. However, given that other details of the results differ between the experiments, a replication would be desirable before firm conclusions are drawn.

3.3. Experiment 3

Experiment 3 tested whether children aged 7–13 years would display weaker evidence than adults of preparatory and selective attention when they were tested on an equivalent task.

3.3.1. Methods

3.3.1.1. Participants

Participants were 26 children (12 male), aged 7–13 years (M = 10.5, SD = 1.7). All participants were declared by their parents to be right-handed native English speakers with no history of hearing problems. All participants had 5-frequency pure-tone average hearing threshold levels of 35 dB or better, tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). Two participants were excluded from the analysis—one due to a technical error during data collection and another due to low behavioural performance in Location trials during the Test Condition (20.8%). It was evident that the child had forgotten the association between the location cues and the target talker. The study was approved by the Research Ethics Committee of the Department of Psychology, University of York.

3.3.1.2. Stimuli and procedure

Stimuli and procedures were the same as those in Experiment 2, except that children completed only 96 trials in the Control Condition and between 96 and 144 trials in the Test Condition (depending on their level of fatigue). Participants received a short break every 16 trials and a longer break every 48 trials. Before undertaking the Test Condition, children completed 16 familiarisation trials (4 in each attention condition).

3.3.1.3. EEG recording, processing, and analyses

EEG recording, processing, and analyses procedures were the same as those in Experiment 2, with one exception. Due to the higher rate of artifacts in EEG data from children than adults, the artifact rejection criteria were relaxed to maintain a similar proportion of rejected trials as in the adult EEG data (< 12.5%). Since it was estimated that artifacts remained in the data, ICA was applied to the EEG data from every child to correct for remaining eye-blink artifacts.

3.3.2. Results

3.3.2.1. Behavioural results

Conjoint accuracy in identifying the colour and number key words in the Test Condition was moderately high and did not differ between Location (M = 89.4%, SD = 7.46) and Gender (M = 88.6%, SD = 7.98) trials, $t(23) = 0.7$, $p = 0.52$. There were also no significant differences in the accuracy with which the visual cue was identified in the Control Condition between Location (M = 97.5%, SD = 3.25) and Gender (M = 98.0%, SD = 2.08) trials, $t(23) = 0.8$, $p = 0.45$.

3.3.2.2. Event-related potentials

Type-I analyses: Differences between Test and Control Conditions

Location trials

Figure 3.11 illustrates the results of the Type-I analyses. Panels A–D report the analysis that compared ERPs between the Test and Control Conditions on trials in which a Location cue was presented. One significant cluster of activity was identified during the Preparatory Phase (Figure 3.11B) and two significant clusters were identified during the Selective Phase (Figure 3.11C–D; Table 3.1).

Gender trials

Panels E–G of Figure 3.11 illustrate the results of the Type-I analysis that compared ERPs between the Test and Control Conditions on trials in which a Gender cue was presented. No significant clusters of activity were identified during the Preparatory Phase, but two significant clusters were identified during the Selective Phase (Figure 3.11F–G; Table 3.1).

Type-II analyses: Differences between Location and Gender trials

Figure 3.12 illustrates the results of Type-II analyses that compared ERPs between Location and Gender trials in the Test Condition. The analysis identified one significant cluster during the Preparatory Phase (Figure 3.12B) and one significant cluster during the Selective Phase (Figure 3.12C; Table 3.2).

The cluster identified during the Preparatory Phase (Cluster 30) showed a greater difference between Location and Gender trials in the Control Condition ($p < 0.001$; Figure 3.13A), which was demonstrated by a significant interaction between cue type (Location/Gender) and condition (Test/Control) [$F(1,23) = 10.74$, $p = 0.003$; Figure 3.13A]. Therefore, it is not possible to rule out the explanation that the cluster arose from differences in the visual cues, rather than from differences in attentional processes triggered by the cues. When the difference was examined in 50-ms sliding windows, the uncorrected p-value did not fall below 0.05 at any time point over the duration of the cluster (Figure 3.13C).

Location trials

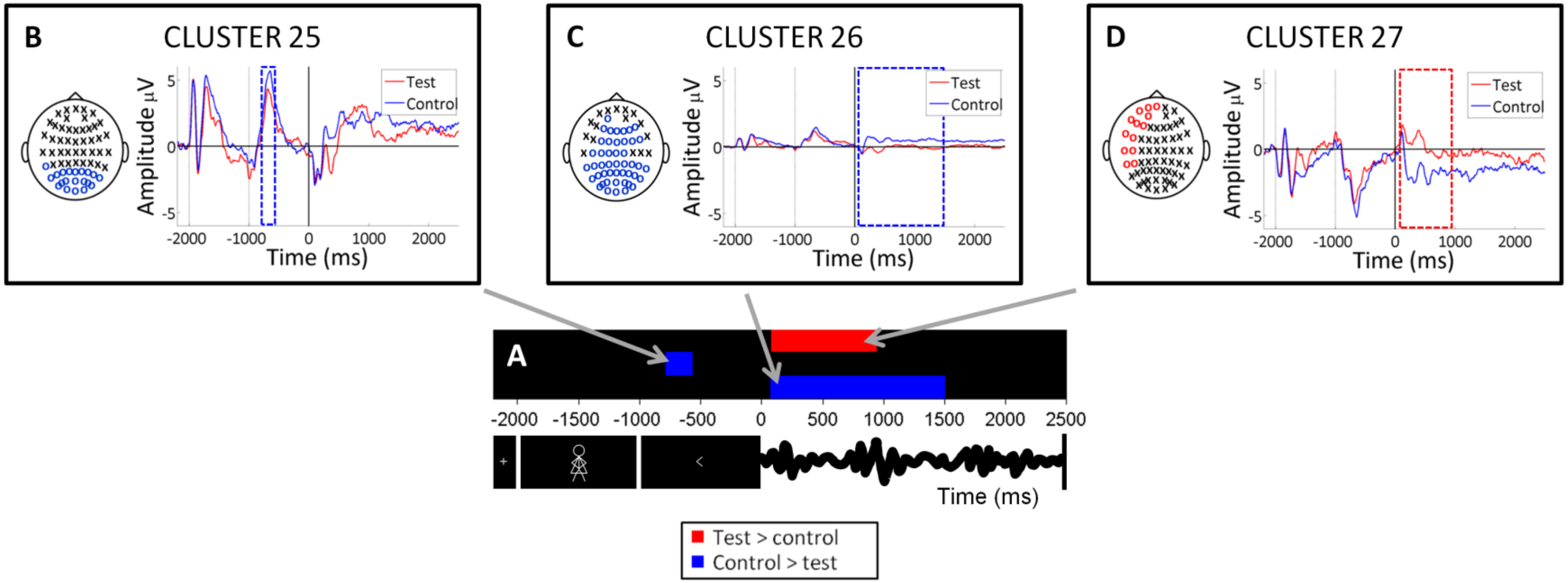


Figure 3.11. (Continued on next page). Experiment 3: Results from Type-I Spatio-temporal Cluster-based Permutation Analyses for Location (**A** to **D**) and Gender (**E** to **G**) trials. (**A** and **E**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Further information about each cluster is displayed in (**B** to **D** and **F** to **G**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

Gender trials

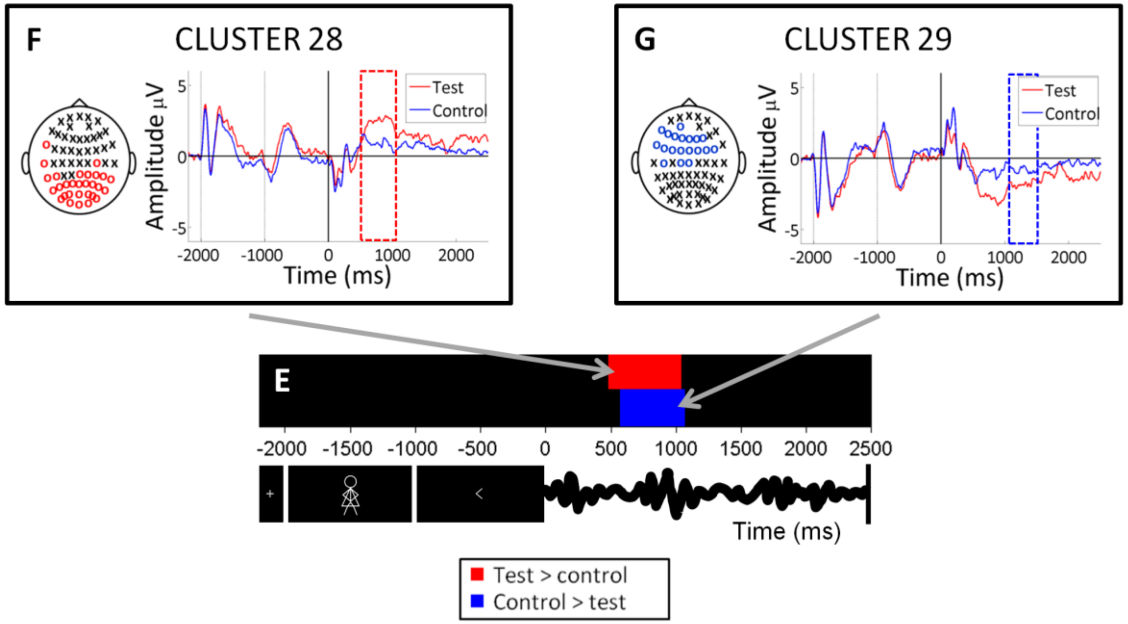


Figure 3.11. (Continued from the previous page)

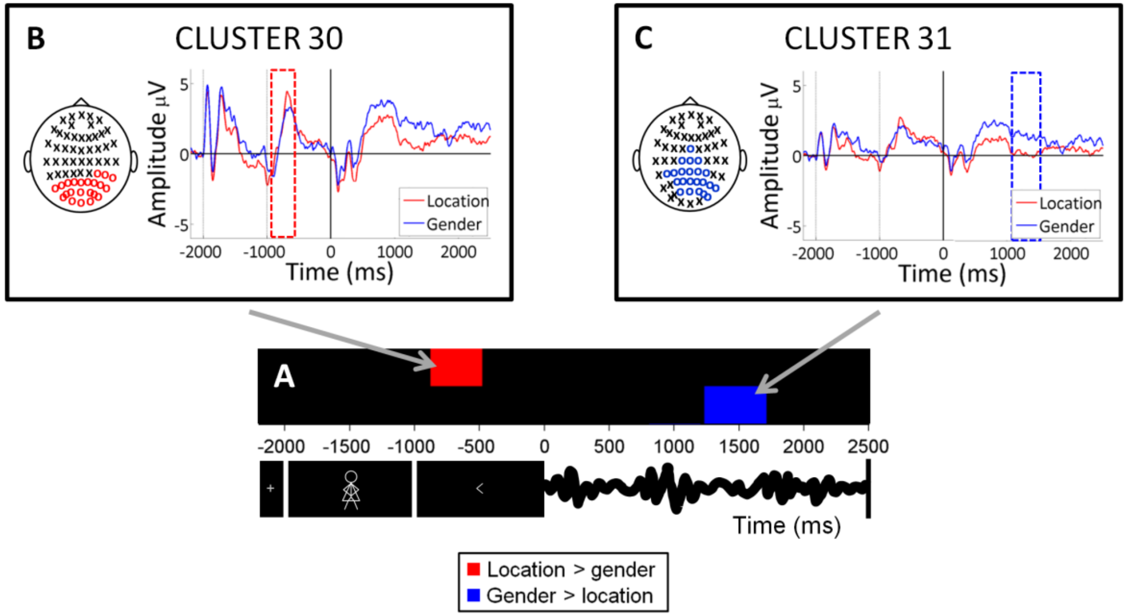


Figure 3.12. Experiment 3: Results from the Type-II Spatio-temporal Cluster-based Permutation Analysis. This analysis contrasted Location and Gender trials in the Test Condition. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive on Location trials than Gender trials. For clusters plotted as blue rectangles, the average amplitude was more negative on Location trials than Gender trials. Further information about each cluster is displayed in (**B**)-(**C**). For each cluster, the topographical map shows the electrodes that contribute to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

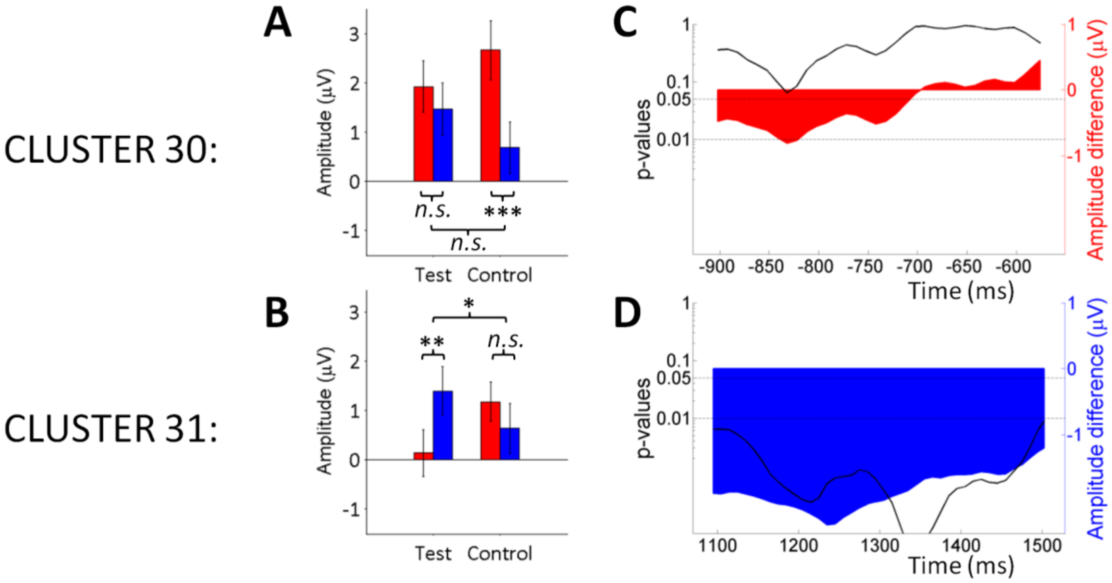


Figure 3.13. Experiment 3: Comparison of differences in the amplitude of ERPs between Location and Gender trials in the Test and Control Conditions for each significant Type-II cluster in Experiment 1. Graphs (**A**)-(**B**) plot the mean amplitude for Location and Gender trials in the Test and Control Conditions, averaged across participants and space-time points. Error bars show 95% within-subjects confidence intervals. Narrow brackets display the significance level of the comparison between Location and Gender trials in the Test and Control Conditions. Wider brackets display the significance level of the two-way interaction (* $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). Graphs (**C**)-(**D**) display the difference of the differences in Gender and Location trials between the Test and Control conditions in 50-ms time windows repeated every 10 ms within the cluster (right axis) and the uncorrected p-values resulting from a paired-samples t-test comparing the differences (left axis). The mid-point of each time window relative to the onset of acoustic stimuli is displayed on the x-axis.

The cluster identified during the Selective Phase (Cluster 31), however, did not show the same pattern in the Control Condition (p = 0.07; Figure 3.13B). Furthermore, the interaction between cue type (Location/Gender) and condition (Test/Control) was significant [$F(1,23) = 13.19$, $p = 0.001$; Figure 3.13B]. This finding demonstrates that the cluster during the Selective Phase arose from differences in the processes for attending selectively to a talker between Location and Gender trials. Figure 3.13D shows that the p-values remained below the 0.01 (uncorrected) level throughout the entire cluster.

3.3.3. Discussion

Similar patterns of ERPs arose in Experiment 3 as in Experiments 1 and 2. Experiment 3 showed significant activity during the Preparatory Phase of Location trials (Figure 3.11A). Type-II differences between Location and Gender trials during the Preparatory Phase were present both in the Test and in the Control Conditions (Figures 3.12B and 3.13A). During the Selective Phase, there were Type-I differences in both Location and Gender trials (3.10A and E, respectively). Type-II clusters during the Selective Phase revealed more negative amplitudes on Location than Gender trials at central electrodes, which were not present in the Control Condition.

However, one key difference between the results of Experiments 2 and 3 was the absence of significant activity during the Preparatory Phase for Gender trials in Experiment 3 (Figure 3.11E). Behavioural accuracy was high and suggests that the children: (1) understood what the cues meant; and (2) were able to select the correct talker based on the gender information provided. Instead, the absence of a significant difference during the Preparatory Phase of Gender trials might have arisen from the fact that children only completed 16 familiarisation trials, due to time constraints, whereas the adults completed 52 familiarisation trials. It is possible that 16 trials were not sufficient for the children to learn the talker characteristics of the male and female talkers, which may have led to a similar problem as Experiment 1—if participants had not learnt the F0 associated with the talkers used in this experiment, then they may not have been able to make specific predictions during the Preparatory Phase, leading to an absence of preparatory brain activity. A possible alternative explanation is that the data in Experiment 3 were noisier than in Experiments 1 and 2 and, therefore, Experiment 3 did not have sufficient power to detect significant clusters during the Preparatory Phase of Gender trials.

3.4. General discussion

All three experiments revealed preparatory EEG activity when participants were cued to the location of a target talker. This result was demonstrated by significant differences in ERPs between the Test and Control Conditions (Figures 3.5, 3.8, and 3.11), despite the fact that the stimuli were identical up to and including the Preparatory Phase for each trial. Therefore, the activity could be attributed to attentional preparation for the upcoming task of selecting one of the two talkers. In Experiments 1 and 2, preparatory activity for location occurred within 50 ms of the full reveal of the visual cue (Clusters 1 and 14; Figures 3.5B and 3.8B) and was sustained for longer than 600 ms during the 1000-ms Preparatory Phase. This result demonstrates that adults begin to prepare their attention early after a location cue is revealed and utilise preparatory brain activity for a large portion of the available time.

Another consistent finding was significant differences between Location and Gender trials during the Selective Phase of the Test Condition. It was possible to rule out the alternative explanation that physical aspects of the visual stimuli were responsible for these differences. Therefore, this differential activity is likely to reflect differences in the mechanisms that participants use to pick out a talker based on their location or gender while the talkers are speaking. Differential activity between Location and Gender trials started with latencies longer than 350 ms after the onset of acoustical stimuli and lasted up to 1500 ms (Figures 3.6, 3.9, and 3.12). In these experiments, the first portion of the sentence did not contain key words that participants had to report. The key words occurred towards the end of the sentence and the long latency of ERPs might reflect this aspect of the stimuli and task.

Accompanying these effects, some additional aspects of activity were likely to result from differences in the visual and acoustical stimuli that were presented in different conditions. A consistent difference between Location and Gender trials was observed early during the Preparatory Phase of the Test Condition (Figures 3.6, 3.9, and 3.12) and similar amplitudes occurred in the Control Condition (Figures 3.7, 3.10, and 3.13). This result likely reflects differences in physical attributes of the visual cues between Location and Gender trials, such as luminance, structural complexity, or differences in the cognitive processes evoked by animate (human stick figures) and inanimate (chevron) cues. Another consistent difference occurred between the Test and Control Conditions during the Selective Phase. The acoustical stimuli in the Control Condition were designed to have the same overall energy and gross fluctuations in amplitude as the pairs of sentences in the Test Condition. However, eliminating cues for location and gender meant that the acoustical stimuli differed in spectral detail, temporal fine structure, and inter-aural differences in level and timing. The Type-I clusters observed during the Selective Phase were sustained throughout most of that phase and were broadly similar in Location and Gender trials (Figures 3.5, 3.8, and 3.11). This observation is consistent with the hypothesis that differences between the Test and Control Conditions during the Selective Phase resulted from differences in the acoustical stimuli.

3.4.1. Preparation and selection by location or gender

Preparatory activity was found consistently when participants were cued to the location of a talker, but not when participants were cued to gender. A Type-I cluster can be significant in one type of trial (Location or Gender), but not the other, for either of two reasons. First, there may be a genuine difference in brain activity in one type of trial but not the other. Alternatively, there may be a difference in brain activity in both conditions but it falls short of significance in one of the conditions. In order to test whether the cluster that occurred during the Preparatory Phase of Location trials also occurred during Gender trials, average amplitudes were compared directly between Location and Gender Conditions. First,

Table 3.3. Summary of comparisons between the Type-I clusters identified during the Preparatory Phase of Location or Gender trials ('Condition in which the cluster occurred' row) and amplitudes—averaged over the electrodes and time points that contributed to the cluster—for the opposite condition (listed in bold font in the first row). For Experiments 1 and 3, clusters that occurred during Location trials were analysed for Gender trials. For Experiment 2, the cluster that occurred during Location trials was analysed for Gender trials and the cluster that occurred during Location trials was analysed for Gender trials. The 'Cluster Number' row shows the cluster that was tested. The 'Polarity' row indicates the condition in which more positive amplitudes were observed, on average across the electrodes and time points at which the cluster occurred for the condition and experiment listed in the first row of the table. A tick in the column headed 'Significant difference between Test/Control Conditions?' indicates that a paired-samples *t*-test revealed a significant difference in amplitude between the Test and Control Conditions in the condition and listed in bold at the top row of the column (*p*-values displayed underneath). A tick in the column headed 'Significant interaction between hearing groups?' indicates that a 2 x 2 between-subjects ANOVA with the factors cue type (Location/Gender) and condition (Test/Control) revealed a significant two-way interaction.

	Experiment 1 Gender	**Experiment 2 Location**	**Experiment 2 Gender**	**Experiment 3 Gender**
Cluster Number	**1**	**17**	**14**	**25**
Condition in which the cluster occurred	Experiment 1 Location	Experiment 2 Gender	Experiment 2 Location	Experiment 3 Location
Significant difference between Test/Control Conditions?	✗ ($p > 0.99$)	✓ ($p = 0.003$)	✓ ($p = 0.013$)	✓ ($p = 0.013$)
Polarity of difference between Test/Control	Control > Test	Control > Test	Control > Test	Test > Control♦
Significant interaction?	✓ ($p = 0.035$)	✗ ($p = 0.26$)	✗ ($p = 0.30$)	✓ ($p = 0.001$)

♦ Different direction to the condition in the first row.

the average amplitude—averaged across the electrodes and time points of the significant cluster during the Preparatory Phase of Location trials—was compared between the Test and Control Conditions in Gender trials in a paired-samples t-test. Second, the amplitudes were compared in a 2 x 2 ANOVA with the factors cue type (Location/Gender) and condition (Test/Control). The results are listed in Table 3.3.

For the clusters identified during Location trials of Experiments 1 and 3, there was no significant difference between the Test and Control Conditions for Gender trials. Furthermore, there was a significant two-way interaction. These results demonstrate that participants in Experiments 1 and 3 used the Preparatory Phase to prepare their attention for the location of an upcoming talker, but not for their gender.

In Experiment 2, significant clusters were observed during the Preparatory Phases of Location *and* Gender trials. A similar analysis was used to identify whether the clusters identified in Location and Gender trials were different (Table 3.3). The pattern of amplitudes were similar for Location and Gender trials. (When the cluster that occurred during Location trials was tested for Gender trials, there was a significant difference between the Test and Control Conditions, and *vice versa*; in addition, the two-way interaction was not significant for either comparison). This finding demonstrates that participants in Experiment 2 displayed similar neural activity when they prepared for the location and gender of an upcoming talker.

Together, the results of Experiments 1 and 2 suggest that adults are able to prepare their attention for an upcoming talker based on a cue for gender, but only if the specific talker is known in advance. When there were several possible male and female talkers to which the gender cue could refer (Experiment 1), participants did not show significant preparation before the talkers started speaking. However, presenting the same identities for the male and female talkers in Experiment 2 might have provided participants with the opportunity to prepare their attention for the F0 and vocal tract length of the male and female talkers. The finding that participants prepare for a talker based on cues that indicate the identity of a male or female talker is compatible with a previous experiment that showed more accurate speech intelligibility when participants knew the identity of an upcoming target talker than when they did not (Kitterick et al., 2010). The idea that differences in ERPs between Experiments 1 and 2 result from differences in the specificity of evidence for gender is consistent with previous experiments that show that the amplitude of ERPs are affected by the predictability of a cue (e.g. Horvath, Sussman, Winkler, & Schröger, 2011; Sussman, Winkler, & Schröger, 2003).

3.4.2. Domain-general and cue-specific effects

The finding that ERPs were similar during the Preparatory Phases of Location and Gender trials in Experiment 2 provides evidence for domain-general preparatory attention. The finding of domain-general activity is consistent with the fMRI results reported by Hill and Miller (2010). They reported overlapping activity in a left-dominant fronto-parietal network in response to a visual cue for location or F0, before three talkers started speaking.

The comparison between Location and Gender trials aimed to reveal whether there was additional cue-specific processing, as reported by Hill and Miller (2010) and Lee *et al.* (2013). In Experiments 1, 2, and 3, cue-specific activity during the Preparatory Phase could be explained by differences in the visual cues, rather than differences in attentional processing of the cues (Table 3.2). Although the visual cues presented by Hill and Miller and

Lee *et al.* had higher similarity than the cues presented in these experiments, the experiments of Hill and Miller and Lee *et al.* did not attempt to rule out the explanation that differences in the visual cues led to differences in brain activity. Therefore, it is possible that activity reported in the experiments of Hill and Miller and Lee *et al.* reflect a combination of activity evoked by physical aspects of the visual cues and preparatory attention for location and F0. In the current experiments, a stringent comparison was performed for cue-specific activity. This comparison had the potential to rule out the alternative explanation that EEG activity was evoked by physical differences between cues for location and gender, but did so at the expense of detecting subtle cue-specific differences in attentional processing.

During the Selective Phase, however, there was evidence for consistent cue-specific activity that could not be explained by differences in the visual cues (Table 3.2). This finding is consistent with the results of Hill and Miller (2010) and Lee *et al.* (2013), who both found significant differences in brain activity when participants selectively attended to a talker, depending on whether participants received information about the talker's spatial location or their F0.

3.4.3. Differences between adults and children

Overall, there are extensive similarities between the results of Experiments 2 and 3, for example, the timing of significant differences and the scalp locations at which significant differences occurred (Table 3.1). This finding is compatible with the finding that children, like adults, can achieve higher accuracy of speech intelligibility from advance cueing in noisy listening environments (Dhamani et al., 2013). The results of Experiment 3 extend the results of Dhamani *et al.* by showing that children aged 7–13 years utilise some of the same brain activity as adults during multi-talker listening.

In order to check whether the results reported in Experiment 3 were consistent throughout the age range tested, or whether significant effects were representative of only a sub-set of the children tested, correlations were performed between age and the amplitude of ERPs within significant clusters or between age and accuracy. The results of the correlations are reported in Appendix A. There were no significant correlations between age and the amplitude of ERPs or between age and accuracy. Therefore, the results reported in Experiment 3 are likely to be consistent across the age range tested.

One main difference in the clusters observed between Experiment 3 and Experiments 1 and 2 was the finding of fewer significant clusters in Experiment 3. In addition, significant clusters in Experiment 3 tended to have shorter durations than significant clusters in Experiments 1 and 2 (Table 3.1). This result is consistent with the hypothesis that children display weaker evidence of preparatory and selective attention than adults. Although, a possible alternative explanation for fewer significant clusters in children is that ERPs recorded from children were more variable than from adults—possibly due to increased noise resulting from the fact that fewer trials contributed to the average waveform for each participant.

3.4.4. Outstanding questions

Overall, there was no evidence of cue-specific activity that was likely to reflect differences in preparatory attention based on cues for location or gender. One possible explanation is that the experiment was too easy for normally-hearing adults. Experiments 1 and 2 showed average accuracy above 90%. Therefore, it is possible that participants would have received no benefit from preparing their attention before the talkers started speaking.

There are two key aspects of the design that might have contributed to high accuracy. First, when presented with two talkers who each spoke two key words, it is likely that participants could monitor both the target and non-target talkers and retain all of the key words in memory. Second, the acoustic stimuli were on average 2.5 seconds long and the key words, whose identities participants had to report, occurred towards the end of each sentence. Therefore, participants had approximately 1.5 seconds during the acoustical stimuli to direct their attention to the target talker before the key words began. Consequently, preparing attention during the Preparatory Phase might not have been necessary or advantageous for the accuracy of speech intelligibility in this task.

The experiments reported in the next chapter investigated whether participants showed higher accuracy of speech intelligibility when the visual cue was presented in advance of the acoustical stimuli (as in the experiments reported in the current chapter) compared to when the cue was presented at the same time as the talkers started speaking. This manipulation addressed the question of whether the Preparatory Phase was necessary for accurate speech intelligibility on this task.

3.4.5. Conclusions

In summary, Experiments 1, 2, and 3 showed consistent evidence of preparatory brain activity when adults and children aged 7–13 years were cued to the location of an upcoming talker. Preparatory activity in adults started early (< 50 ms) after a visual cue for location was fully revealed and was sustained for longer than 600 ms. Taken together, the results of Experiments 1 and 2 suggest that adults display preparatory brain activity when they know the gender of an upcoming talker, but only when the gender cue predicts the identity of the target talker. Experiments 1, 2, and 3 also provide evidence for cue-specific

EEG activity when two talkers spoke simultaneously—the amplitude of brain activity depended on whether participants attended selectively to a talker based on information about their location or gender.

Chapter 4

The Effect of Preparation Time on Speech Intelligibility

The two experiments reported in this chapter aimed to explore the behavioural benefit of advance cueing for speech intelligibility during multi-talker listening. Although previous research has consistently demonstrated that knowing characteristics of an upcoming talker improves intelligibility (e.g. Best, Ozmeral, et al., 2007; Kidd et al., 2005; Kitterick et al., 2010), little is known about the time course of preparatory attention. When cueing the location or identity of an upcoming talker, previous experiments have tended to use different cue-target intervals, ranging between 100 ms before the target (Koch et al., 2011) to cueing at the beginning of each block of trials (Brungart & Simpson, 2007; Ericson et al., 2004; Kitterick et al., 2010). No similar experiments have systematically varied the cue-target interval within a single experiment.

With respect to the length of the cue-target interval, there are at least two possibilities: (1) the length of the cue-target interval does not improve intelligibility until it reaches a threshold, beyond which longer intervals do not improve intelligibility further; or (2) longer cue-target intervals improve intelligibility progressively. These experiments aimed to distinguish between these alternatives by measuring the accuracy and latency of speech intelligibility when the duration of the cue-target interval was varied between 0 ms and 2000 ms. The cue-target interval was measured as the duration of time between the presentation of a visual cue for location or gender and the onset of two or three talkers.

Experiment 1 used a task in which two talkers spoke simultaneously. Based on the experiments reported in Chapter 3, which found high overall accuracy in a two-talker listening task, varying the duration of the cue-target interval in Experiment 1 was expected to have no effect on the accuracy or latency of speech intelligibility. Experiment 2 used a more difficult task in which three talkers spoke simultaneously. For Experiment 2, cue-target intervals greater than 0 ms (i.e. when the cue was presented before the talkers started speaking) were expected to lead to higher accuracy and shorter latencies than the 0-ms interval (i.e. when the cue was revealed at the same time that the talkers started speaking), since previous experiments have shown improved speech intelligibility when participants know the spatial location or identity of a talker before he or she begins to speak (e.g. Best, Marrone, & Mason, 2007; Kitterick et al., 2010). However, there were no prior expectations about how accuracy and latency would vary between the shortest and longest cue-target interval.

4.1. Experiment 1

4.1.1. Methods

4.1.1.1. Participants

Participants were 20 young adults (10 male), aged 18–27 years (mean [M] = 20.1, standard deviation [SD] = 2.1). They were self-declared native English speakers with no history of hearing problems. The study was approved by the Research Ethics Committee of the Department of Psychology of the University of York.

4.1.1.2. Apparatus and stimuli

Apparatus and stimuli (Figures 4.1 and 4.2) were the same as those used in the Test Condition of Experiment 2 reported in Chapter 3.

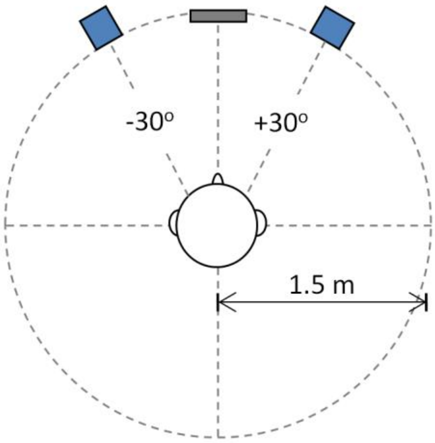


Figure 4.1. Layout of loudspeakers (blue squares) and visual display unit (grey rectangle) relative to a participant's head in Experiment 1.

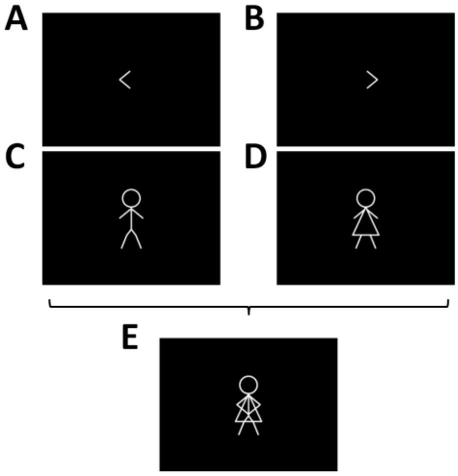


Figure 4.2. (**A**)-(**D**) Visual cues. (**E**) Visual composite stimulus, which is a combination of the four visual cues overlaid.

Figure 4.3. Trial structure of Experiment 1, with an example trial below.

4.1.1.3. Procedure

At the start of each trial, a fixation cross was presented for 1000 ms (Figure 4.3). Next, the visual composite stimulus was presented, which faded to reveal the visual cue for each trial. The fade lasted 200 ms. The total amount of time between the onset of the visual composite stimulus and the onset of the acoustical stimuli was fixed at 3000 ms. Although, the duration of the visual cue varied quasi-randomly from trial to trial. There were five possible intervals between the full reveal of the visual cue and the onset of the acoustical stimuli: 0, 250, 500, 1000, and 2000 ms. Two different sentences were presented from the two loudspeakers (left and right). The sentences started simultaneously, but contained different call signs and different colour-number combinations. One talker was male and the other was female. The two talker identities remained the same over the course of the experiment and the male and female talkers were presented equally often from each location.

The visual cue directed attention to the target talker and varied quasi-randomly from trial to trial. The cue remained on the screen throughout the duration of the acoustic stimuli so that participants did not have to retain the visual cue in memory. Participants were instructed to report the colour-number combination in the target sentence by pressing a coloured digit on a touch screen directly in front of their chair. They were instructed to respond as quickly and as accurately as possible. The coloured digits appeared on the screen before each trial and participants were able to respond at any point during the trial. Participants were instructed to look at the central video screen, although their heads were not restrained. The inter-trial interval varied randomly from 1000 to 1500 ms. Each participant completed 360 trials (72 for each cue duration and, within this, 18 trials for each of the different visual cues), with a break every 40 trials.

The logic behind the design was that, on every trial, there was a fixed time interval (3000 ms) between the onset of the visual composite stimulus and the onset of the acoustical stimuli. This aspect ensured that any differences between different cue-target intervals must be explained by differences in the duration of time in which participants received information about the location or gender of the upcoming talker. Any advantage for longer cue-target intervals, therefore, could not be explained by a general increase in arousal for longer cue-target intervals or by changes in the predictability of the onset time of the acoustical stimuli.

Prior to the main task, participants completed two sets of familiarisation trials. In the first set, 12 trials were presented in which *either* the male or female talker was presented on each trial from the left or right loudspeaker. The aim was to familiarise participants with the left and right locations and with the male and female talkers that would be used in the main task. The trial structure was the same as the main task but only one or other of the two talkers, but not both, was presented on each trial. The second set of familiarisation trials were identical to the main task. Participants completed 4 trials (1 for each visual cue). Each trial contained both voices. During both sets of familiarisation trials, the cue-target interval varied randomly from trial-to-trial.

4.1.1.4. Analyses

Trials were separated into attend-location (average left/right cues) and attend-gender (average male/female cues) groups, separately for each of the five cue-target interval conditions.

For each trial, three categories of response were recorded: (1) correct identification of both the colour and number (i.e. the “Colour-number combination”) spoken by the target talker; (2) correct identification of the colour irrespective of whether the number was reported correctly (“Colour-only”); (3) correct identification of the number irrespective of whether the colour was reported correctly (“Number-only”). In addition, reaction times (RTs), measured from the onset of the acoustical stimuli, were calculated on trials in which participants correctly identified the Colour-number combination. RTs beyond two standard deviations from the mean for each participant were excluded from the analysis.

4.1.2. Results

4.1.2.1. Colour-number accuracy

Colour-number accuracy was high across all cue-target intervals (Figure 4.4A). A 5 x 2 repeated-measures ANOVA was conducted with the factors cue-target interval (5 levels) and cue type (location/gender). There was a main effect of cue-target interval, $F(4, 76) = 3.23$, $p = 0.017$, $\eta_p^2 = 0.15$. Contrasts showed that the 500-ms interval led to significantly lower Colour-number accuracy than the 0-ms cue-target interval [$F(1, 19) = 4.58$, $p = 0.046$, $\eta_p^2 = 0.19$]. However, none of the other cue-target intervals had Colour-number accuracies significantly different to the 0-ms cue-target interval.

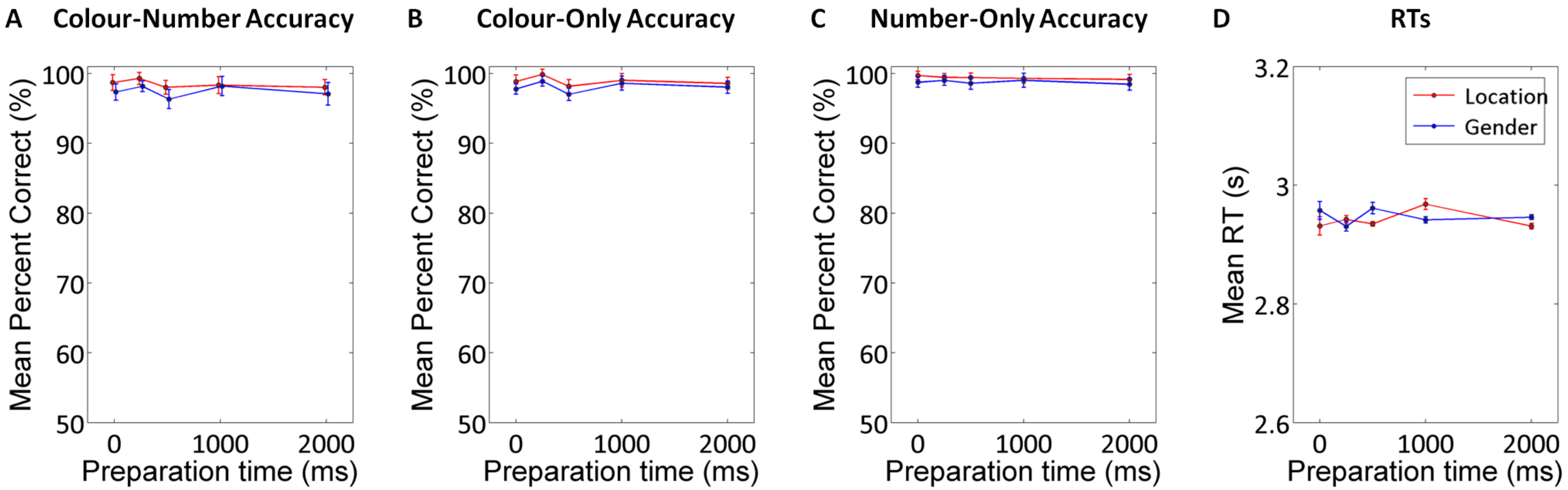


Figure 4.4. Accuracy and reaction time (RT) results for Experiment 1. (**A**) Mean percentage of trials where participants correctly identified the colour-number combination spoken by the target talker. (**B**) Mean percentage of trials where participants correctly identified the colour (irrespective of the number). (**C**) Mean percentage of trials where participants correctly identified the number (irrespective of the colour). (**D**) Mean RTs on trials in which participants correctly identified the colour-number combination. Error bars show within-subjects 95% confidence intervals.

There was no significant difference in Colour-number accuracy between attend-location and attend-gender trials (Figure 4.4A), $F(1, 19) = 3.64$, $p = 0.07$. There was also no significant interaction between cue-target interval and cue type [$F(2.1, 40.8) = 0.57$, $p = 0.58$].

4.1.2.2. Colour-only accuracy

Colour-only accuracy was also high across all cue-target intervals (Figure 4.4B). There was a significant main effect of cue-target interval, $F(3.0, 56.6) = 4.75$, $p = 0.005$, $\eta_p^2 = 0.20$. Contrasts showed that the 500-ms interval led to significantly lower Colour-only accuracy than the 0-ms cue-target interval [$F(1, 19) = 6.69$, $p = 0.018$, $\eta_p^2 = 0.26$]. However, none of the other cue-target intervals had colour accuracies significantly different to the 0-ms cue-target interval.

There was no significant effect of cue type [$F(1, 19) = 3.66$, $p = 0.07$] and no significant interaction between cue-target interval and cue type [$F(2.0, 38.1) = 0.23$, $p = 0.80$].

4.1.2.3. Number-only accuracy

There was no significant effect of cue-target interval on Number-only accuracy (Figure 4.4C), $F(2.9, 55.7) = 0.56$, $p = 0.64$. In addition, there was no significant main effect of cue type [$F(1, 19) = 3.1$, $p = 0.10$] and no significant interaction between cue-target interval and cue type [$F(2.9, 54.2) = 0.40$, $p = 0.74$].

4.1.2.4. RTs

Figure 4.4D displays the RTs for each cue-target interval, measured from the onset of the acoustical stimuli. There was a significant main effect of cue-target interval, $F(2.0, 37.9) = 10.47$, $p < 0.001$, $\eta_p^2 = 0.36$. Contrasts showed that 250-ms [$F(1, 19) = 7.00$, $p = 0.016$, $\eta_p^2 = 0.27$] and 2000-ms [$F(1, 19) = 17.14$, $p = 0.001$, $\eta_p^2 = 0.47$] led to significantly shorter RTs than the 0-ms interval. The 1000-ms interval led to significantly longer RTs than the 0-ms interval [$F(1, 19) = 7.57$, $p = 0.013$, $\eta_p^2 = 0.29$] and the 500-ms interval did not show significantly different RTs to the 0-ms interval [$F(1, 19) = 2.27$, $p = 0.15$].

There was no significant difference in RTs when participants knew the location of the target talker (M = 2.94 s, SD = 0.08), compared to when they knew the gender of the target talker (M = 2.95 s, SD = 0.07), $F(1, 19) = 3.15$, $p = 0.09$. However, there was a significant two-way interaction between cue-target interval and cue type, $F(1.4, 26.2) = 11.30$, $p = 0.001$, $\eta_p^2 = 0.37$.

4.1.3. Discussion

Accuracy in reporting key words spoken by the target talker was high across Colour-number, Colour-only, and Number-only accuracy. Colour-number and Colour-only accuracy showed significant main effects of cue-target interval, although this effect was driven by lower accuracy for the 500-ms cue-target interval compared to the 0-ms interval, which was the opposite direction to the prediction.

The effect of cue-target interval on RTs was unclear. There was only slight variation in RTs across different cue-target intervals (average RTs varied only by 30 ms); although, due to small within-subjects confidence intervals, some of the differences between the cue-target intervals were statistically significant. Overall, there was no systematic effect of shorter compared to longer cue-target intervals on the length of RTs.

One factor that might have contributed to high accuracy across all cue-target intervals was the long duration (2.5 seconds) of the acoustic stimuli. Furthermore, the colour and number key words, whose identities participants had to report, occurred towards the end of each sentence. Therefore, participants had approximately 1.5 seconds during the acoustical stimuli to direct their attention to the target talker before the key words began and, consequently, they may not have needed to utilise the cue-target interval to prepare their attention.

4.2. Experiment 2

Experiment 2 presented a more difficult task to obtain accuracy below ceiling level. Three modifications were applied to the task: (1) loudspeaker positions were fixed at ± 15° azimuth to reduce spatial separation; (2) a third 'distracting' talker was presented from a central loudspeaker (0° azimuth); and (3) the duration of the acoustical stimuli were shortened, such that the colour and number key words were spoken at the beginning of each sentence.

The first modification was intended to increase the difficulty of talker segregation, since previous experiments report smaller spatial release from masking with smaller degrees of spatial separation (Marrone et al., 2008c; Noble & Perrett, 2002). The second modification was intended to increase perceptual load. Previous experiments have revealed decreased accuracy for speech intelligibility when the number of competing talkers is increased from one to three (Hawley, Litovsky, & Culling, 2004). In addition, Ericson *et al.* (2004) found a significant improvement in the accuracy of reporting words spoken by a target talker when participants received information about the location the upcoming target talker, but only for three-talker listening and not for two-talker listening. The third modification decreased the amount of time that participants could prepare for the target talker during the acoustical stimuli, meaning that the cue-target interval was a more valid indicator of the amount of time that participants could use to prepare their attention for the upcoming talker. One aim of

Experiment 2 was to investigate whether the duration of the cue-target interval affected the accuracy and latency of speech intelligibility during multi-talker listening.

A second aim was to explore the types of errors that participants made when they did not correctly identify words spoken by the target talker. Previous studies have reported that errors during multi-talker listening typically consist of words spoken by competing talker(s), rather than words that were not spoken on that trial (Brungart & Simpson, 2002b; Darwin, 2006). During three-talker listening, it was possible to take this idea further in order to gain insights into whether participants had correctly segregated the three talkers or not. If the talkers had been segregated successfully, then errors would be expected to consist of key words that were spoken by only one of the presented talkers on that trial. However, if the talkers had not been segregated effectively, then errors would be as likely to consist of words spoken by a mixture of the talkers as words spoken by only one talker.

A third aim of Experiment 2 was to explore whether participants were attending to the location and gender of a talker in combination, or to only the cued attribute on each trial. The logic of this analysis arose from the well-established switch cost effect—the finding that RTs are longer when participants have to switch attention to a different attribute than when participants maintain attention on the same attribute (S Monsell & Driver, 2000; Rogers & Monsell, 1995). This analysis focussed on trials in which participants received the same visual cue on two consecutive trials and compared trials in which the non-cued attribute remained the same as the previous trial with trials in which the non-cued attribute changed.

4.2.1. Methods

4.2.1.1. Participants

Participants were 20 young adults (10 male), aged 18–24 years (M = 19.6, standard SD = 1.8). They were self-declared native English speakers with no history of hearing problems. The study was approved by the Research Ethics Committee of the Department of Psychology of the University of York.

4.2.1.2. Apparatus

Apparatus were the same as Experiment 1 with the exception that loudspeakers (Plus XS.2, Canton) were located at 0° (fixation) and at 15° to the left and right (Figure 4.5).

4.2.1.3. Stimuli

Visual stimuli were identical to Experiment 1, including the visual cues (“left”, “right”, “male”, and “female”).

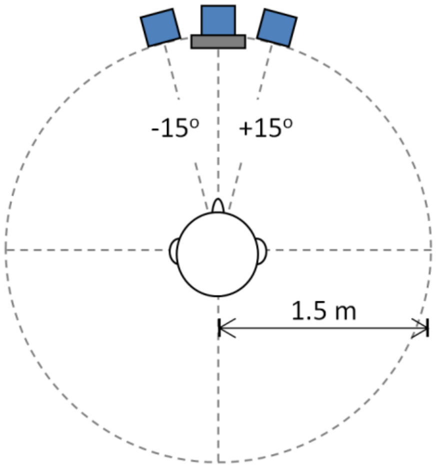


Figure 4.5. Layout of loudspeakers (blue squares) and visual display unit (grey rectangle) relative to a participant's head in Experiment 2.

The acoustical stimuli were derived from the stimuli presented in Experiment 1. The original stimuli were edited so that each sentence had the form '<colour> <number> now'. There were four colours ('blue', 'red', 'green', 'white') and four numbers ('1', '2', '3', '4'). An example is "Green Two now". The identity of the male and female voices remained the same as Experiment 1. An additional female talker was selected from the corpus, whose voice was manipulated to sound like a "child's" voice by simulating a change in F0 and vocal tract length using Praat (Version 5.3.08; http://www.praat.org/). The average duration of the new sentences was 1.4 s. The levels of the digital recordings of the sentences were normalised to the same root mean square (RMS) power.

4.2.1.4. Procedure

The trial structure was the same as that used in Experiment 1 (Figure 4.6) except that the composition of acoustical stimuli differed. One sentence was played from each loudspeaker (left, centre, and right) with the same onset time but a different colour-number combination. The "child" voice was always played from the central loudspeaker and was never the target. Of the remaining two voices, one was always the male and the other was always the female and they were presented equally often at the left and right loudspeakers.

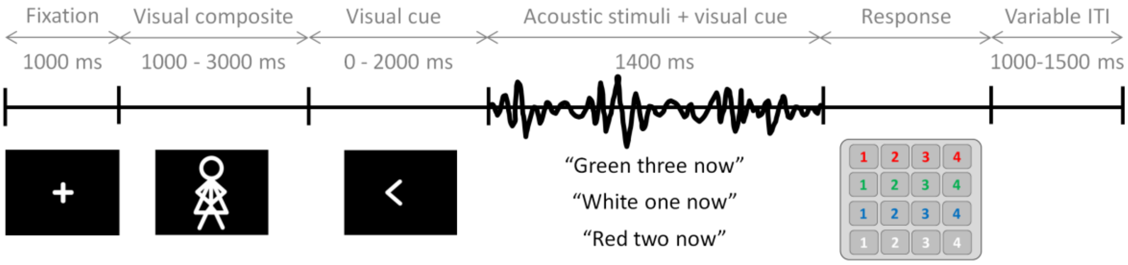


Figure 4.6. Trial structure of Experiment 2, with an example trial below.

Each participant completed 360 trials (72 for each cue duration and, within this, 18 trials for each of the different visual cues), with a break every 40 trials.

Prior to the main task, participants completed two sets of familiarisation trials with the same structure as Experiment 1. In the first set (12 trials), *either* the male or female talker was presented on each trial from the left or right loudspeaker. In the second set (4 trials), each trial contained all three voices.

4.2.1.5. Analyses

Trials were separated into attend-location (average left/right cues) and attend-gender (average male/female cues) groups, separately for each of the five preparation time conditions.

Accuracy and RTs

For each trial, three categories of response were recorded: (1) correct identification of both the colour and number (i.e. the "Colour-number combination") spoken by the target talker; (2) correct identification of the colour irrespective of whether the number was reported correctly ("Colour-only"); (3) correct identification of the number irrespective of whether the colour was reported correctly ("Number-only"). In addition, RTs, measured from the onset of the acoustical stimuli, were calculated on trials in which participants correctly identified the Colour-number combination. RTs beyond two standard deviations from the mean for each participant were excluded from the analysis.

Errors

When participants did not correctly identify the Colour-number combination, responses were categorised into one of four different types of error. The reported Colour-number combination could be: (1) spoken by the opposite-gender talker that was presented from the contralateral location ("opposite-gender" error), (2) spoken by the "child" talker that was presented from the central location ("child" error), (3) a mixture of words spoken by the target and a non-target talker *or* a mixture of words spoken by the two non-target talkers ("mix" error), or (4) not be spoken by any mixture of the talkers on that trial ("absent" error).

The percentages of the four types of error were assessed in relation to the percentages expected if participants guessed randomly with a uniform distribution. The expected percentages were: 6.7% "opposite-gender" error, 6.7% "child" error, 40.0% "mix" error, and 46.7% "absent" error.

Trial-by-trial analysis

A trial-by-trial analysis was used to determine whether participants were using 'object-based' attention (i.e. attending to the location and gender of a talker simultaneously) or location- and feature-based attention. This novel analysis was inspired by the 'switch cost' in performance when participants have to change their attentional focus from one feature to another (S Monsell & Driver, 2000; Rogers & Monsell, 1995). Trials were included in this analysis only if the cue for that trial was identical to the previous trial. Colour-number accuracy and RTs were compared between trials in which the array of talkers had the same compared to a different configuration to the previous (n-1$^{\text{th}}$) trial. For the attend-gender condition, trials in which the target talker was the same gender and had the same location as the previous trial were compared with trials in which the target talker was the same gender but had a different location (Figure 4.7A–B). For the attend-location condition, trials in which the target talker was the same gender and had the same location as the previous trial were compared with trials in which the target talker was the opposite gender but had the same location (Figure 4.7C–D).

The rationale behind this novel approach was that accuracy and RTs for 'object-based' attention would be influenced by the non-cued dimension. Based on this hypothesis, accuracy would be lower and RTs longer when the array of talkers had a different configuration to the previous trial than when the talkers had the same configuration (i.e. displaying a 'switch

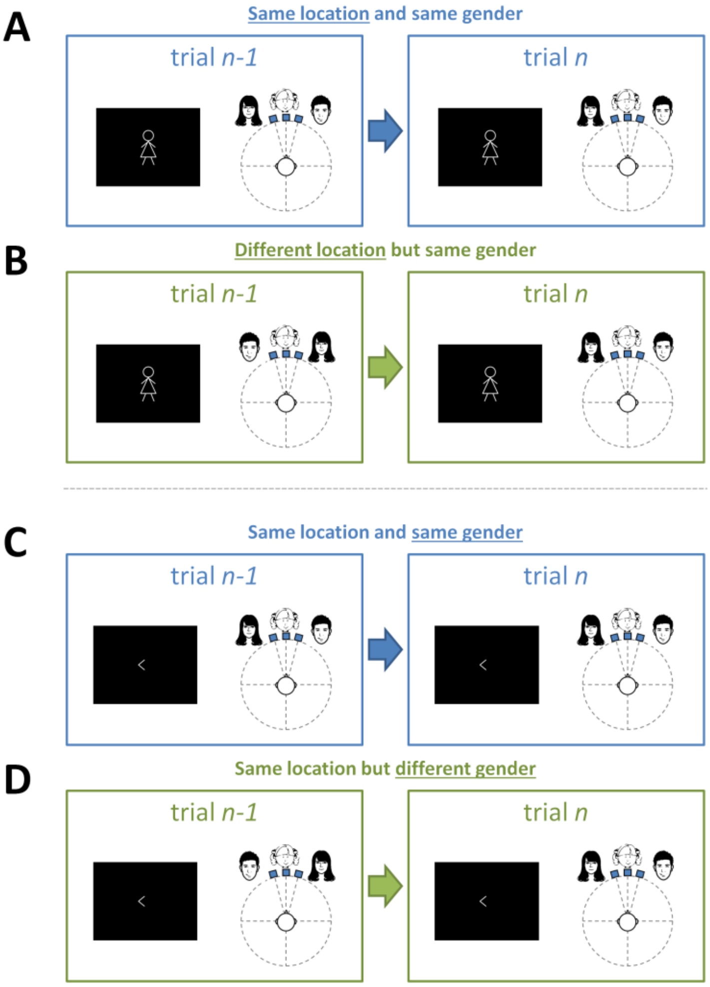


Figure 4.7. Schematic of trial-by-trial analysis displaying example trials. For attend-gender trials (the example displays the "female" cue), the analysis compared colour-number accuracy and reaction times (RTs) in trials where the target talker had the same location and gender as the previous (n-1$^{\text{th}}$) trial (**A**) to trials where the target talker had the same gender but a different location to the previous trial (**B**). Panels (**C**)-(**D**) show the equivalent comparison for attend-location trials (the example displays the "left" cue).

cost'), even though the visual cue was identical to the previous trial. However, if participants were using 'location-based' or 'feature-based' attention in attend-location and attend-gender trials respectively, accuracy and RTs should not be affected by the configuration of talkers when the visual cue was identical to the previous trial.

4.2.1.6. Colour-Number Accuracy

Figure 4.8A illustrates the results for Colour-number accuracy. A 5 x 2 repeated-measures ANOVA showed a significant main effect of cue-target interval, $F(2.9, 54.1) = 3.50$, $p = 0.023$, $\eta_p^2 = 0.16$. Contrasts showed that 500-ms [$F(1, 19) = 8.71$, $p = 0.008$, $\eta_p^2 = 0.31$] and 2000-ms [$F(1, 19) = 22.49$, $p < 0.001$, $\eta_p^2 = 0.54$] intervals led to significantly higher Colour-number accuracy than the 0-ms cue-target interval. Neither of the other cue-target intervals had Colour-number accuracies that were significantly higher than the 0-ms interval.

Participants achieved higher Colour-number accuracy in the attend-location condition (M = 87.8%, SD = 4.7) than the attend-gender condition (M = 84.2%, SD = 5.2), $F(1, 19) = 13.75$, $p = 0.001$, $\eta_p^2 = 0.42$. There was no significant interaction between cue-target interval and cue type [$F(4, 76) = 0.24$, $p = 0.92$].

4.2.1.7. Colour-only Accuracy

The pattern for Colour-only accuracy (Figure 4.8B) was similar to the pattern observed for Colour-number accuracy. There was a significant main effect of cue-target interval, $F(2.8, 52.9) = 3.47$, $p = 0.025$, $\eta_p^2 = 0.16$. Similarly, 500-ms [$F(1, 19) = 8.29$, $p = 0.010$, $\eta_p^2 = 0.30$] and 2000-ms [$F(1, 19) = 17.66$, $p < 0.001$, $\eta_p^2 = 0.48$] intervals led to significantly higher Colour-only accuracy than the 0-ms interval. Neither of the other cue-target intervals had colour accuracies significantly higher than the 0-ms interval.

Colour-only accuracy was higher in the attend-location condition (M = 88.5%, SD = 4.4) than the attend-gender condition (M = 85.9%, SD = 4.6), $F(1, 19) = 7.87$, $p = 0.011$, $\eta_p^2 = 0.29$. There was no significant interaction between cue-target interval and cue type [$F(4, 76) = 0.48$, $p = 0.75$].

4.2.1.8. Number-only Accuracy

There was no significant effect of cue-target interval on Number-only accuracy (Figure 4.8C; $F(4, 76) = 1.37$, $p = 0.25$]. However, Number-only accuracy was significantly higher in the attend-location condition (M = 98.5%, SD = 1.7) than the attend-gender condition (M = 97.1%, SD = 2.0), $F(1, 19) = 15.13$, $p = 0.001$, $\eta_p^2 = 0.44$. There was no significant interaction between cue-target interval and cue type [$F(2.8, 53.8) = 0.41$, $p = 0.74$].

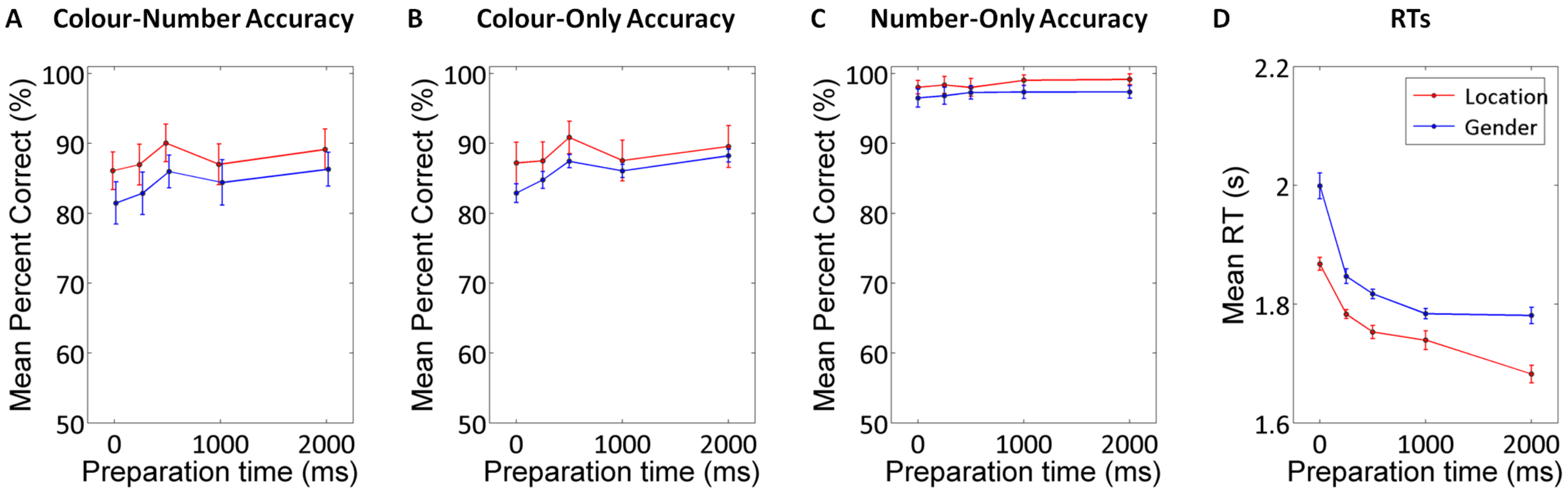


Figure 4.8. Accuracy and reaction time (RT) results for Experiment 2. (**A**) Mean percentage of trials where participants correctly identified the colour-number combination spoken by the target talker. (**B**) Mean percentage of trials where participants correctly identified the colour (irrespective of the number). (**C**) Mean percentage of trials where participants correctly identified the number (irrespective of the colour). (**D**) Mean RTs on trials in which participants correctly identified the colour-number combination. Error bars show within-subjects 95% confidence intervals.

4.2.1.1. RTs

RTs became shorter as the duration of the cue-target interval increased (Figure 4.8D). There was a main effect of cue-target interval, $F(1.4, 27.4) = 213.40$, $p < 0.001$, $\eta_p^2 = 0.92$. Contrasts showed that the 250-ms [$F(1, 19) = 590.86$, $p < 0.001$, $\eta_p^2 = 0.97$], 500-ms [$F(1, 19) = 442.39$, $p = 0.001$, $\eta_p^2 = 0.96$], 1000-ms [$F(1, 19) = 297.37$, $p < 0.001$, $\eta_p^2 = 0.94$], and 2000-ms [$F(1, 19) = 283.25$, $p < 0.001$, $\eta_p^2 = 0.94$] intervals produced significantly shorter RTs than the 0-ms interval. Bonferroni-corrected post-hoc tests also showed significant differences in RTs between all adjacent preparation times ($p \leq 0.001$).

RTs were significantly shorter in the attend-location condition (M = 1.8 s, SD < 0.1) than the attend-gender condition (M = 1.9 s, SD < 0.1), $F(1, 19) = 461.39$, $p < 0.001$, $\eta_p^2 = 0.96$. There was also a significant two-way interaction between cue-target interval and cue type, $F(2.0, 38.4) = 103.13$, $p < 0.001$, $\eta_p^2 = 0.84$.

4.2.1.2. Errors

The largest percentage of errors were “mix” errors (M = 78.4%, SD = 9.1), where the reported Colour-number combination was spoken by a mixture of the presented talkers. The second largest percentage of errors were “absent” errors (M = 17.2%, SD = 8.7), where the colour and/or number was not spoken by any of the talkers on that trial. Participants made “opposite-gender” errors (M = 3.5%, SD = 4.3) and “child” errors (M = 1.0%, SD = 1.3) on a low proportion of trials. The percentages of “mix” [$t(19) = 19.33$, $p < 0.001$] and “absent” [$t(19) = 15.70$, $p < 0.001$] errors were significantly greater than their expected values, whereas the percentages of “opposite-gender” [$t(19) = 3.99$, $p = 0.001$] and “child” [$t(19) = 24.25$, $p < 0.001$] errors were significantly smaller than their expected values.

A 4 x 5 x 4 repeated-measures ANOVA investigated whether the types of errors (4 levels: “opposite-gender”, “child”, “mix”, and “absent” errors) differed significantly between the different cue-target intervals (5 levels: 0, 250, 500, 1000, and 2000 ms) or between cue types (4 levels: left, right, male, and female). There was a significant main effect of error type, $F(1.5, 27.61) = 367.20$, $p < 0.001$, $\eta_p^2 = 0.95$. Bonferroni-corrected post-hoc tests showed that the percentage of “opposite-gender” errors did not differ significantly from the percentage of “child” errors ($p = 0.18$), but there were significant differences between the percentages of all other error type combinations ($p < 0.001$).

There was no significant difference in the percentages of errors made for different cue-target intervals [error type * cue-target interval interaction: $F(4.5, 84.5) = 0.57$, $p = 0.71$] and no significant difference in the percentages of errors made across the four different cue types [error type * cue type interaction: $F(3.1, 59.8) = 1.09$, $p = 0.36$].

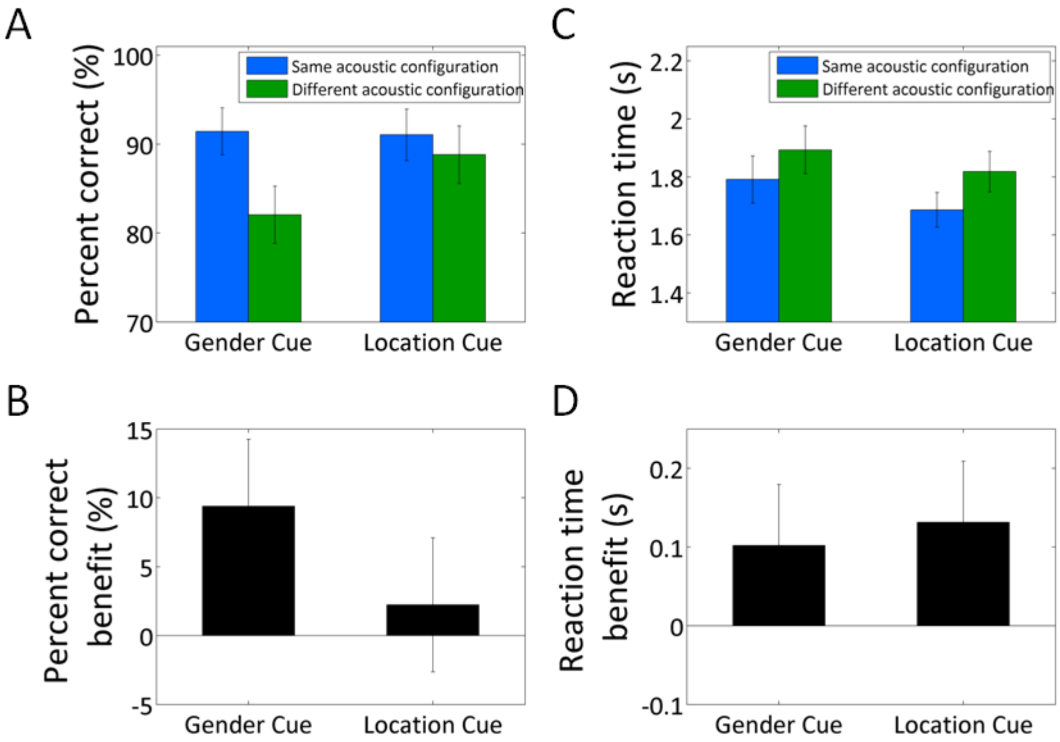


Figure 4.9. Results from the trial-by-trial analysis. (**A**) Accuracy for reporting the colour-number combination spoken by the target talker, separated by cue type (Location/Gender), when the acoustic configuration was either the same (i.e. the location and gender of the target talker was the same) or different (i.e. the target talker varied on the uncued dimension) to the previous trial. (**B**) Accuracy benefit, calculated as the difference in percent correct when the acoustic configuration was the same as the previous trial compared to when it was different. (**C**)-(**D**) show equivalent plots for RTs. Error bars display within-subjects 95% confidence intervals.

4.2.1.3. Trial-by-trial analysis

The trial-by-trial analysis revealed higher Colour-number accuracy (Figure 4.9A) and shorter RTs (Figure 4.9C) when the configuration of talkers was the same as the previous trial than when it was different. Separate 2 x 2 ANOVAs were performed on the Colour-number accuracy and RT data, with the factors of configuration (same/different) and cue type (location/ gender).

Colour-number accuracy was significantly higher when participants were cued to location than gender, which is consistent with the results reported above (Section 4.2.2.1), $F(1, 19) = 4.94$, $p = 0.039$, $\eta_p^2 = 0.21$. Trials with the same configuration as the previous trial (M = 91.2%, SD = 3.9) displayed significantly higher accuracy than trials with a different configuration (M = 85.4%, SD = 6.1), $F(1, 19) = 23.4$, $p < 0.001$, $\eta_p^2 = 0.55$. There was also a significant two-way interaction, with gender trials leading to a larger difference in accuracy between the same and different configuration conditions than location trials (Figure 4.9B), $F(1, 19) = 4.75$, $p = 0.042$, $\eta_p^2 = 0.20$.

RTs were significantly shorter when participants were cued to location than gender, which is consistent with the results reported above (Section 4.2.2.4), $F(1, 19) = 22.88$, $p < 0.001$, $\eta_p^2 = 0.52$. There was also a main effect of configuration, with same-configuration trials (M = 1.7 s, SD = 0.2) displaying significantly shorter RTs than different-configuration trials (M = 1.9 s, SD = 0.2), $F(1, 19) = 20.58$, $p < 0.001$, $\eta_p^2 =0.52$. The interaction was not significant (Figure 4.9D; $F(1, 19) = 0.32$, $p = 0.58$).

4.2.2. Results

4.2.3. Discussion

The average accuracy with which participants reported the Colour-number combination in Experiment 2 was lower than Experiment 1. This result demonstrates that the task used in Experiment 2 was more difficult than used in Experiment 1.

Presenting the cue before the onset of the acoustical stimuli led to higher Colour-number and Colour-only accuracy than presenting the cue at the onset of the acoustical stimuli. Specifically, 500-ms and 2000-ms intervals led to significantly higher accuracy than the 0-ms interval. There was no significant difference between accuracy at the 250-ms and 0-ms cue-target intervals. It is possible, therefore, that a 250-ms interval did not provide enough time for participants to successfully prepare for an upcoming talker. It is unclear why there was no significant benefit of a 1000-ms interval over a 0-ms interval, since a benefit was observed at both shorter (500-ms) and longer (2000-ms) intervals. One possibility is that this effect was obscured by a speed-accuracy trade-off, since RTs were significantly shorter for the 1000-ms than the 0-ms interval.

RTs became significantly shorter as the cue-target interval increased from 0-ms to 2000-ms. This result is consistent with the explanation that increasing the amount of preparation time improves performance. Rather than showing a threshold for successful preparation, increasing the amount of preparation time progressively improved RTs—each cue-target interval led to significantly shorter RTs than the previous cue-target interval. The progressive improvement of RTs with increasing durations of preparation time was present in all participants (Appendix E).

The RT data showed a significant interaction between the direction of attention and the cue-target interval. This result shows that increasing the amount of preparation time did not affect both cue types in the same manner. The significant interaction appears to be largely driven by the shortest and longest cue-target intervals. The difference in RT between 250-ms and 0-ms was greater in attend-gender than attend-location trials. This finding is consistent with the idea that preparation provides a greater benefit when baseline RTs are longer.

4.3. General discussion

During two-talker listening (Experiment 1), participants achieved high accuracy of speech intelligibility even when the cue was fully revealed at the same time as the talkers started speaking. Therefore, the results did not show systematic effects of the duration of the cue-target interval on the accuracy or latency of speech intelligibility. However, for the three-talker listening task (Experiment 2), for which average accuracy was lower, RTs systematically shortened as the duration of the cue-target interval increased.

Previous experiments in which advance cueing was compared to no advance cueing have demonstrated a behavioural advantage from knowing the spatial location (Best, Gallun, Carlile, & Shinn-Cunningham, 2007; Best, Ihlefeld, & Shinn-Cunningham, 2005; Ericson et al., 2004; Kidd et al., 2005) or the identity (Kitterick et al., 2010) of the target talker before he or she begins to speak. For example, Ericson *et al.* (2004) found a significant improvement in the accuracy of reporting words spoken by a target talker when participants received information about the location of the upcoming target talker. However, Ericson *et al.* only found this result for three-talker listening and not for two-talker listening—in their two-talker condition, accuracy was near-ceiling even when participants did not receive a cue. Similar results have also been reported by Brungart *et al.* (2001). The different pattern of results between Experiments 1 and 2 are consistent with the results of Ericson *et al.* and Brungart *et al.*, although, since the duration of the acoustical stimuli also differed between Experiments 1 and 2, it is possible that at least some of the differences observed between the Experiments 1 and 2 could be attributable to this aspect.

The results of Experiment 2 build upon the results of previous experiments by showing that the *duration* of the cue-target interval affects the accuracy and latency of speech intelligibility. One previous experiment varied the interval between the onset of an instructive cue and the onset of an acoustical target stimulus (Richards & Neff, 2004)—although, in contrast to the current experiments, participants had to detect the presence or absence of a pure tone among a masking complex tone. Richards and Neff found that thresholds for detecting pure tones were worse for a 5-ms cue-target interval than for intervals of 50, 100, 250, and 500 ms. However, there were no significant differences between any of the other intervals. This result suggests that participants gain some benefit from having 50-ms to prepare for the target, but no further improvement between 50-ms and 500-ms. The results of Experiment 2 extend these findings by showing that RTs shorten progressively when participants have longer than 500 ms of preparation time.

4.3.1. Colour-only and Number-only accuracy

In both experiments, accuracy for reporting the number only (irrespective of colour) was high and did not differ significantly across cue-target intervals. In contrast, the pattern for Colour-only accuracy was similar to the accuracy of reporting the Colour-number combination. In combination, these findings suggest that the majority of errors were due to incorrect identification of the colour rather than the number key word. One possible explanation is that participants need time to 'tune in' to the talkers during the presentation of acoustical stimuli, since the colour key word always preceded the number key word. However, a possible alternative explanation is that the number key words were more distinguishable from each other than the colour key words, irrespective of the time at which they were spoken.

4.3.2. Incorrect responses

The possible origin of incorrect responses was inferred from the data in Experiment 2. The highest proportion of errors consisted of mixtures of words spoken by different talkers (which were either combinations of the target and non-target talkers or combinations of the two non-target talkers). This result is consistent with the explanation that participants failed to segregate the talkers on incorrect trials (i.e. difficulties with 'object formation').

4.3.3. Attention to the task-irrelevant dimension

The trial-by-trial analysis provided evidence that participants attended to both the location and the gender of the target talker, which is consistent with the idea of 'object-based' attention. On trials in which the visual cue was identical to the previous trial, RTs were shorter when the configuration of talkers remained the same as the previous trial compared to when the configuration changed from the previous trial. For example, on consecutive trials in which the participant was cued to the female talker, RTs were shorter when the female talker was on the left on both trials than when the female talker was on the left on one trial and the right on the next trial. This finding demonstrates that a task-irrelevant attribute can influence the accuracy of speech intelligibility, which suggests that participants attended to spatial and non-spatial attributes of the talker in combination during this task. The accuracy data were consistent with the idea that the RT results could not be explained by a speed-accuracy trade-off.

Importantly, the same pattern of RTs were observed on attend-location and attend-gender trials, which is inconsistent with the alternative explanation that participants were *either* using space-based *or* feature-based attention on both types of trial. If participants were directing space-based attention during *both* attend-location and attend-gender trials, then RTs should be affected by the location of the talker on attend-gender trials, but RTs should *not* be affected by the gender of the talker on attend-location trials.

One possible reason why participants may have adopted attention to both location and gender in this task is that the acoustical stimuli were natural speech, which fluctuates over time. When identifying words spoken during multi-talker listening in everyday life, it would be advantageous to monitor multiple cues at once rather than focusing on a single cue. The dynamic nature of speech signals means that the factors most useful for segregating talkers vary over time (Caporello Bluvas & Gentner, 2013). Within-talker F0 fluctuations or talkers who are moving whilst speaking are both factors that could contribute to differences in the cues that are most useful at any point in time.

4.3.4. Outstanding questions

These results have implications for the interpretation of the experiments reported in Chapter 3, in which brain activity was measured during a two-talker listening task. The finding that there was no systematic improvement in the accuracy or latency of speech intelligibility with increasing cue-target intervals in the two-talker task used in Experiment 1 is consistent with the idea that it was not necessary for participants to engage preparatory attention before the talkers began. Therefore, it is possible that significant brain activity reported in Chapter 3 underestimates the amount of preparatory brain activity that would be observed in a more challenging task, in which participants achieve better speech intelligibility when they have time to prepare their attention before a target talker starts speaking. The three-talker task employed in Experiment 2 showed a systematic effect of increasing the duration of the cue-target interval on the latency of speech intelligibility. In addition, accuracy was higher for the 2000-ms and 500-ms cue-target intervals than the 0-ms interval. Therefore, the three-talker task was expected to show a greater extent of preparatory EEG activity than the results reported for the two-talker task described in Chapter 3. The experiment reported in the next chapter measured brain activity in a three-talker task that was the same as Experiment 2, except that the duration of the cue-target interval was fixed at 2000 ms on every trial.

4.3.5. Conclusions

In summary, these experiments provide evidence that longer durations of preparation time lead to higher accuracy and shorter latencies for reporting words spoken by a target talker in thee-talker listening. The results, however, do not provide evidence for a benefit of longer preparation times in two-talker listening. In the two-talker task, accuracy was high even when participants had no time to prepare before the talkers started speaking.

The results of Experiment 2 distinguish two alternative explanations by which preparation time was hypothesised to influence speech intelligibility. Rather than a ‘threshold’ amount of time for successful preparation, the results showed that increasing the duration of preparation time progressively improved the latency with which participants correctly reported target words.

Chapter 5

Brain Activity during Three-talker Listening

This experiment aimed to investigate the time course of brain activity during three-talker listening. Of the two previous experiments that have investigated brain activity in preparation for multi-talker listening (Hill & Miller, 2010; Lee et al., 2013), only one investigated brain activity during three-talker listening (Hill & Miller, 2010). Hill and Miller (2010) measured brain activity using functional magnetic resonance imaging (fMRI) and found preparatory brain activity in a left-hemisphere fronto-parietal network. In addition, the detailed pattern of activity within the network depended on whether participants were preparing to select the upcoming target talker based on their spatial location or fundamental frequency (F0). The current experiment measured brain activity using electro-encephalography (EEG), with the aim of revealing preparatory brain activity with higher temporal resolution than the previous experiment using fMRI.

This experiment also aimed to build upon the results reported in Chapter 3, which measured the time course of brain activity during *two-talker* listening. The results reported in Chapter 4 imply that preparatory brain activity was not necessary or beneficial for speech intelligibility in the two-talker task employed in Chapter 3. In contrast, the three-talker task employed in Chapter 4 showed an improvement in the accuracy and latency of speech intelligibility when participants had time to prepare for either the location or the gender of an upcoming talker. Based on these findings, the current experiment (which employed a similar three-talker listening task as that used in Chapter 4) was expected to show a greater *extent* of preparatory EEG activity than the experiments reported in Chapter 3. Nevertheless, overall *patterns* of domain-general and cue-specific activity throughout the task were expected to be similar in the current experiment as in the experiments reported in Chapter 3, since aspects of the task design were similar.

Another aim was to estimate the likely neural generators of scalp-recorded EEG activity using minimum norm source reconstruction. Based on the results of Hill and Miller

(2010), it was expected that activity within a fronto-parietal network would underlie significant differences in ERPs during the Preparatory Phase (i.e. between the reveal of a visual cue for location or gender and the onset of the talkers) and activity within a temporo-parietal network would underlie significant differences in ERPs during the Selective Phase (i.e. after the talkers started speaking).

5.1. Methods

5.1.1. Participants

Participants were 20 young adults (7 male), aged 20–31 years (mean [M] = 23.8, standard deviation [SD] = 3.0). They were self-declared right-handed native English speakers with no history of hearing problems. They had 5-frequency average pure-tone hearing levels of 20 dB HL or better, tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). Each participant had previously provided a high-resolution whole-brain structural MRI scan measured with a GE 3 Tesla HDx Excite MRI scanner at the York Neuroimaging Centre. The study was approved by the Research Ethics Committee of the York Neuroimaging Centre of the University of York.

5.1.2. Apparatus

The experiment was conducted in a 5.3 m x 3.7 m single-walled test room (Industrial Acoustics Co., NY) located within a larger sound-treated room. Participants sat facing three loudspeakers (Plus XS.2, Canton) arranged in a circular arc at a height of 1 m at 0° azimuth

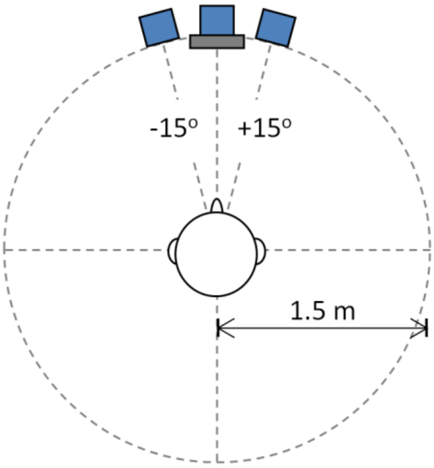


Figure 5.1. Layout of loudspeakers (blue squares) and visual display unit (grey rectangle) relative to a participant's head.

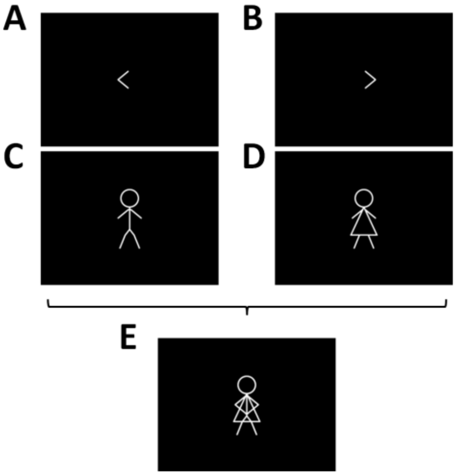


Figure 5.2. (**A**)-(**D**) Visual cues. (**E**) Visual composite stimulus, which is a combination of the four visual cues overlaid.

(fixation) and at 15° to the left and right (Figure 5.1). A 15-inch visual display unit (VDU; NEC AccuSync 52VM) was positioned directly below the central loudspeaker.

5.1.3. Stimuli

5.1.3.1. Visual cues

The visual stimuli were identical to those used in Experiment 2 of Chapter 4 (Figure 5.2). Four visual cues, “left”, “right”, “male”, and “female”, were defined by white lines on a black background. Left and right cues were leftward- and rightward-pointing arrows, respectively; male and female cues were stick figures (Figure 5.2A–D). A composite visual stimulus consisted of the four cues overlaid (Figure 5.2E).

5.1.3.2. Acoustical test stimuli

The acoustical stimuli for the Test Condition were identical to those used in Experiment 2 of Chapter 4.

5.1.3.3. Acoustical control stimuli

Control stimuli were single-channel noise-vocoded representations of concurrent triplets of CRM sentences. Each control stimulus was created by summing a triplet of sentences digitally with their onsets aligned, extracting the temporal envelope of the combination using the Hilbert Transform (Hilbert, 1912), and using the envelope to modulate the amplitude of a random noise whose long-term spectrum matched the average spectrum of all of the possible triplets of sentences.

5.1.4. Procedures

5.1.4.1. Test Condition

Figure 5.3A shows the trial structure for the Test Condition, which was the same as the structure used in Experiment 2 of Chapter 4, except that the duration of the visual cue was fixed at 2000 ms on every trial and the duration of the visual composite was fixed at 1000 ms. Each participant completed 384 trials (96 in each cueing condition), with a break every 48 trials.

5.1.4.2. Control Condition

The trial structure of the Control Condition was the same as the Test Condition (Figure 5.3B) with the exception that an acoustical control stimulus, presented from the loudspeaker at 0° azimuth, replaced the triplet of acoustical test stimuli. The task was to identify the picture that corresponded to the visual cue on each trial. Participants responded by pressing a touch-screen monitor positioned directly in front of their chair. Each participant completed 216 trials (54 in each visual stimulus condition), with a break every 36 trials. The presentation level of the control stimuli was set so that their average level matched the average level of the triplets of test stimuli. Participants undertook the Control Condition before the Test Condition; that is, before they had learnt the association between the visual cues and the acoustical test stimuli.

5.1.4.3. Familiarisation trials

After participants had completed the Control Condition, but before they undertook the Test Condition, they completed two sets of familiarisation trials, which were the same as those in Experiment 2 of Chapter 4. In the first set (12 trials), *either* the male or female talker was presented on each trial from the left or right loudspeaker. In the second set (4 trials), each trial contained all three voices. EEG activity was not recorded during familiarisation.

5.1.5. EEG recording and processing

EEG recording and processing were identical to the experiments reported in Chapter 3.

5.1.6. Behavioural analyses

Trials were separated into Location (average left/right cues) and Gender (average male/female cues) groups, separately for the Test and Control Conditions. Responses were scored as correct if both the colour and number key words were reported correctly in the Test Condition, and if the visual cue was reported correctly in the Control Condition.

5.1.7. Analyses of ERPs

The same types of ERP analyses were conducted as in Chapter 3. Spatio-temporal

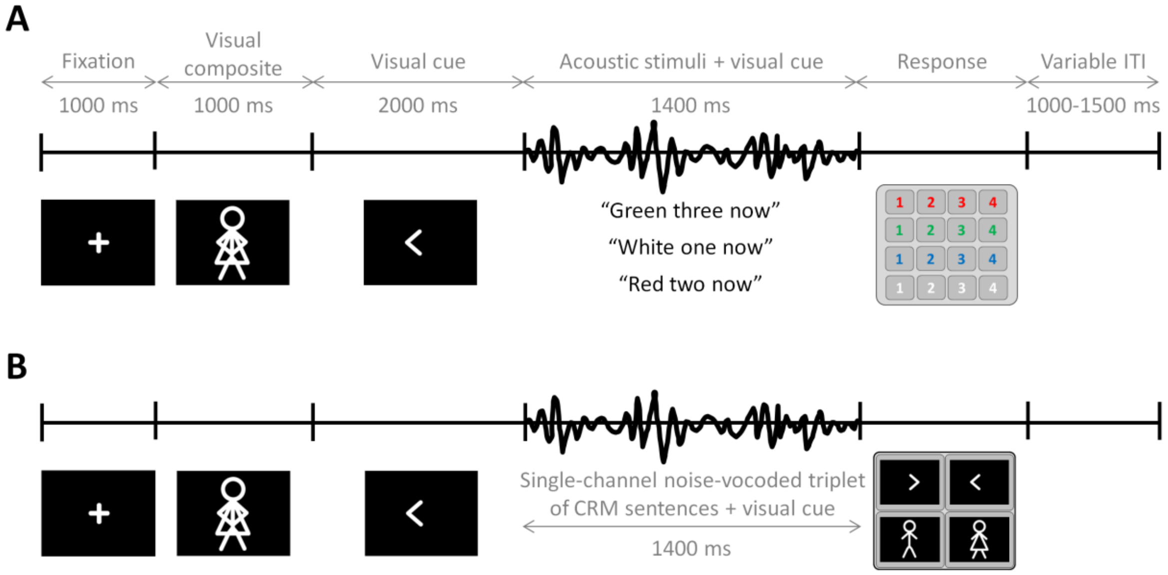


Figure 5.3. (**A**) Trial structure in the Test Condition, with an example trial below. (**B**) Trial structure in the Control Condition.

Cluster-based Permutation Analyses were used to make two types of comparison. In Type-I analyses, the Test and Control conditions were compared, separately for Location and Gender trials. Type-I clusters found in the Preparatory Phase could not arise from sensory or perceptual processes because the stimuli did not differ between the conditions in this phase. Rather, such differences were interpreted as arising from contrasting attentional preparatory activity between the Test and Control conditions. Type-I clusters found in the Selective Phase, in contrast, could arise either from differences in attentional activity or from differences between the acoustical structure of the Test and Control stimuli.

In Type-II analyses, Location with Gender trials were first compared within the Test Condition. These analyses identified clusters where ERPs differed significantly depending on whether participants were receiving cues for, and directing attention towards, location or gender. Such differences could be evoked *either* by different attentional processes or by physical differences between the visual cues. Accordingly, the average amplitude of Location and Gender trials—averaged over the space-by-time points in the cluster—was compared between the Test and Control Conditions in a 2 x 2 repeated-measures ANOVA. A two-way interaction meant that the cluster could not be fully explained by the influence of physical differences in the visual cues between conditions. In order to determine whether such differences were sustained over the entire duration of a cluster or were restricted to particular moments, the difference of the differences in Location and Gender trials between the Test and Control Conditions was plotted, averaging only over the space-by-time points that fell in a 50-ms time window that was advanced in 10-ms steps over the duration of the cluster.

5.1.8. Source reconstruction

Source reconstruction aimed to indicate the location of the neural generators that contributed to scalp-recorded activity. First, the scalp-recorded EEG data was localised to individual voxels in the brain using the SPM8 (Wellcome Trust Centre for Neuroimaging; www.fil.ion.ucl.ac.uk/spm/software/spm8) toolbox for MATLAB. Individual head models were calculated for each participant, which were derived from each participant’s structural MRI scan using voxel sizes of 2 x 2 x 2 mm. Model parameters were estimated using a classical minimum norm model implemented in the SPM8 toolbox (Independent and Identically Distributed error model). Averaged images (1–20 Hz) were created for each participant at 50-ms intervals over the time windows in which significant clusters of ERPs were identified in the Spatio-temporal Cluster-based Permutation Analysis. When analysing the time windows of Type-I clusters, this procedure was conducted separately for Location

Table 5.1. (Continued on next page). Summary of results for the Gender and Location Condition comparisons (Type-I analysis) between the Test and Control Conditions. The row headed 'Cluster p-value' shows the results of the Spatio-temporal Cluster-based Permutation Analyses.

Phase		**Location Condition**	**Gender Condition**
Preparatory	**Cluster Number**	**1**	**8**
	Cluster *p*-value	0.016	0.002
	Polarity	Test > Control	Test > Control
	Electrode Locations	Posterior	Posterior
	Onset of cluster (ms)	322	236
	Duration of cluster (ms)	373	559
Preparatory	**Cluster Number**	**2**	**9**
	Cluster *p*-value	0.032	0.016
	Polarity	Control > Test	Control > Test
	Electrode Locations	Central + Anterior	Central
	Onset of cluster (ms)	24	0
	Duration of cluster (ms)	373	328
Preparatory	**Cluster Number**	**3**	**10**
	Cluster *p*-value	0.022	0.003
	Polarity	Control > Test	Control > Test
	Electrode Locations	Central	Central
	Onset of cluster (ms)	473	360
	Duration of cluster (ms)	289	500
Preparatory	**Cluster Number**	**4**	**11**
	Cluster *p*-value	0.002	0.001
	Polarity	Control > Test	Control > Test
	Electrode Locations	Central	Posterior
	Onset of cluster (ms)	1015	1251
	Duration of cluster (ms)	928	749
Selective	**Cluster Number**	**5**	**12**
	Cluster *p*-value	< 0.001	< 0.001
	Polarity	Test > Control	Test > Control
	Electrode Locations	Non-Central	Non-Central
	Onset of cluster (ms)	0	0
	Duration of cluster (ms)	1200	1200
Selective	**Cluster Number**	**6**	**13**
	Cluster *p*-value	0.001	< 0.001
	Polarity	Control > Test	Control > Test
	Electrode Locations	Central + Posterior	Anterior + Central + Posterior
	Onset of cluster (ms)	0	0
	Duration of cluster (ms)	502	1200

Table 5.1. (Continued from the previous page)

Phase		Location Condition	Gender Condition
Selective	**Cluster Number**	**7**	-
	Cluster *p*-value	0.001	-
	Polarity	Control > Test	-
	Electrode Locations	Anterior	-
	Onset of cluster (ms)	509	-
	Duration of cluster (ms)	691	-

and Gender trials in the Test and Control Conditions. For Type-II clusters, the procedure was conducted separately for Location and Gender trials within the Test Condition.

5.2. Results

5.2.1. Behavioural results

Conjoint accuracy in identifying the colour and number key words in the Test Condition differed significantly between Location (M = 80.6%, SD = 10.7) and Gender (M = 75.8%, SD = 11.6) trials, $t(19) = 4.5$, $p < 0.001$. There were no significant differences in the accuracy with which the visual cue was identified in the Control Condition between Location (M = 99.5%, SD = 0.76) and Gender (M = 99.2%, SD = 0.84) trials, $t(19) = 1.5$, $p = 1.45$.

5.2.2. Event-related potentials

5.2.2.1. Type-I analyses: Differences between Test and Control Conditions

Location trials

Figure 5.4 illustrates the results of the Type-I analyses on trials in which a Location cue (left/right) was presented. During the 2000-ms Preparatory Phase, four significant clusters of activity were identified (Figure 5.4A). Cluster 1 involved 21 posterior electrodes and spanned the time interval from 322 to 695 ms, relative to the start of the phase. Cluster 1 showed significantly more positive amplitude during the Test Condition than the Control Condition (cluster statistic = 10585, $p = 0.016$; Figure 5.4B). The polarity, location, onset time, and duration of Cluster 1 are tabulated in the first line of the first column of Table 5.1. Cluster 2 (Figure 5.4C) spanned the interval from 24 to 397 ms. It involved 34 central and anterior electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 8197, $p = 0.032$]. Cluster 3 (Figure 5.4D; 473 to 762 ms) involved a similar group of electrodes as Cluster 2 but began later in time. Cluster 3 involved 25 central and fronto-central electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 9300, p =

0.022]. Cluster 4 (Figure 5.4E; 1015 to 2000 ms) occurred in the second half of the Preparatory Phase. It involved 30 posterior electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 35082, $p < 0.001$]. Clusters 1–4 demonstrate that differences in brain activity arise between a condition in which a visual cue has no implications for auditory attention and a condition in which the same visual cue directs listeners to prepare to select an upcoming talker on the basis of their location.

During the Selective Phase, three significant clusters of activity were identified (Figure 5.4A). Cluster 5 (Figure 5.4F) spanned the entire Selective Phase (0 to 1200 ms, relative to the start of the phase). Cluster 5 involved 56 electrodes across almost the entire electrode array and showed significantly more positive amplitude during the Test Condition than the Control Condition [cluster statistic = 79846, $p < 0.001$]. Cluster 6 (Figure 5.4G) spanned the interval from 0 to 502 ms, relative to the start of the phase. It involved 39 central and posterior electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 47070, $p = 0.001$]. Cluster 7 (Figure 5.4H; 509 to 1200 ms) occurred shortly after the offset of Cluster 6. Cluster 7 involved 27 anterior electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 44578, $p = 0.001$].

Gender trials

In the second of the Type-I analyses, ERPs in the Test and Control Conditions were compared on trials in which a Gender cue (male/female) was presented. Panels I–P of Figure 5.4 show these results. Four significant clusters were identified during the Preparatory Phase. Cluster 8 involved 17 posterior electrodes and spanned the time interval from 236 to 795 ms, relative to the start of the phase. Cluster 8 showed significantly more positive amplitude during the Test Condition than the Control Condition [cluster statistic = 15904, $p = 0.002$] (Figure 5.4J). Cluster 9 (Figure 5.4K) spanned the interval from 0 to 328 ms. It involved 28 central electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 8543, $p = 0.016$]. Cluster 10 (Figure 5.4L; 360 to 860 ms) involved 36 electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 14482, $p = 0.003$]. Cluster 11 (Figure 5.4M; 1251 to 2000 ms) occurred in the second half of the Preparatory Phase. It involved 27 posterior electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 20083, $p = 0.001$]. Clusters 8–11 provide evidence for differences in brain activity between a

Location trials

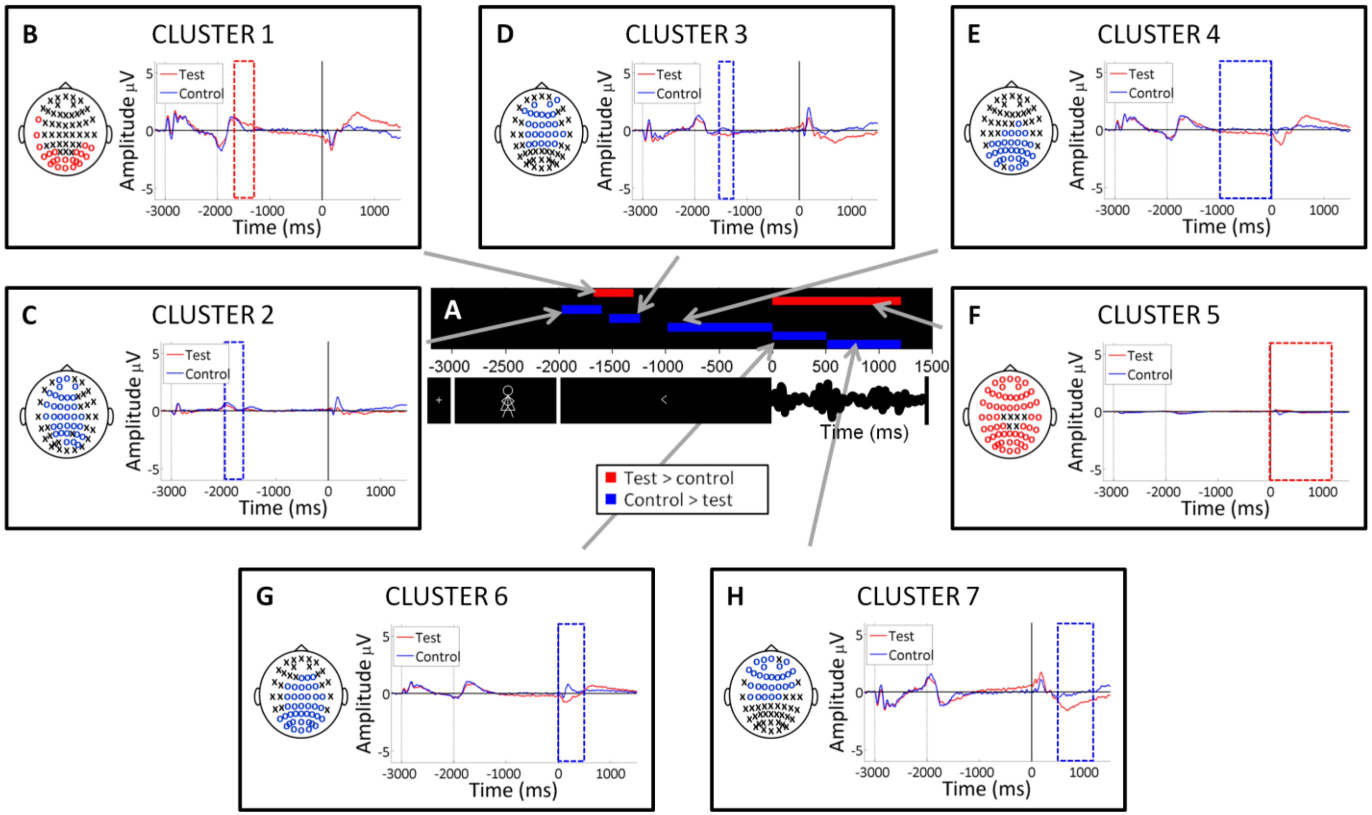


Figure 5.4. (Continued on next page). Results from Type-I Spatio-temporal Cluster-based Permutation Analyses for the Location (**A** to **H**) and Condition (**I** to **P**) trials. (**A** and **I**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Further information about each cluster is displayed in (**B** to **H** and **J** to **P**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

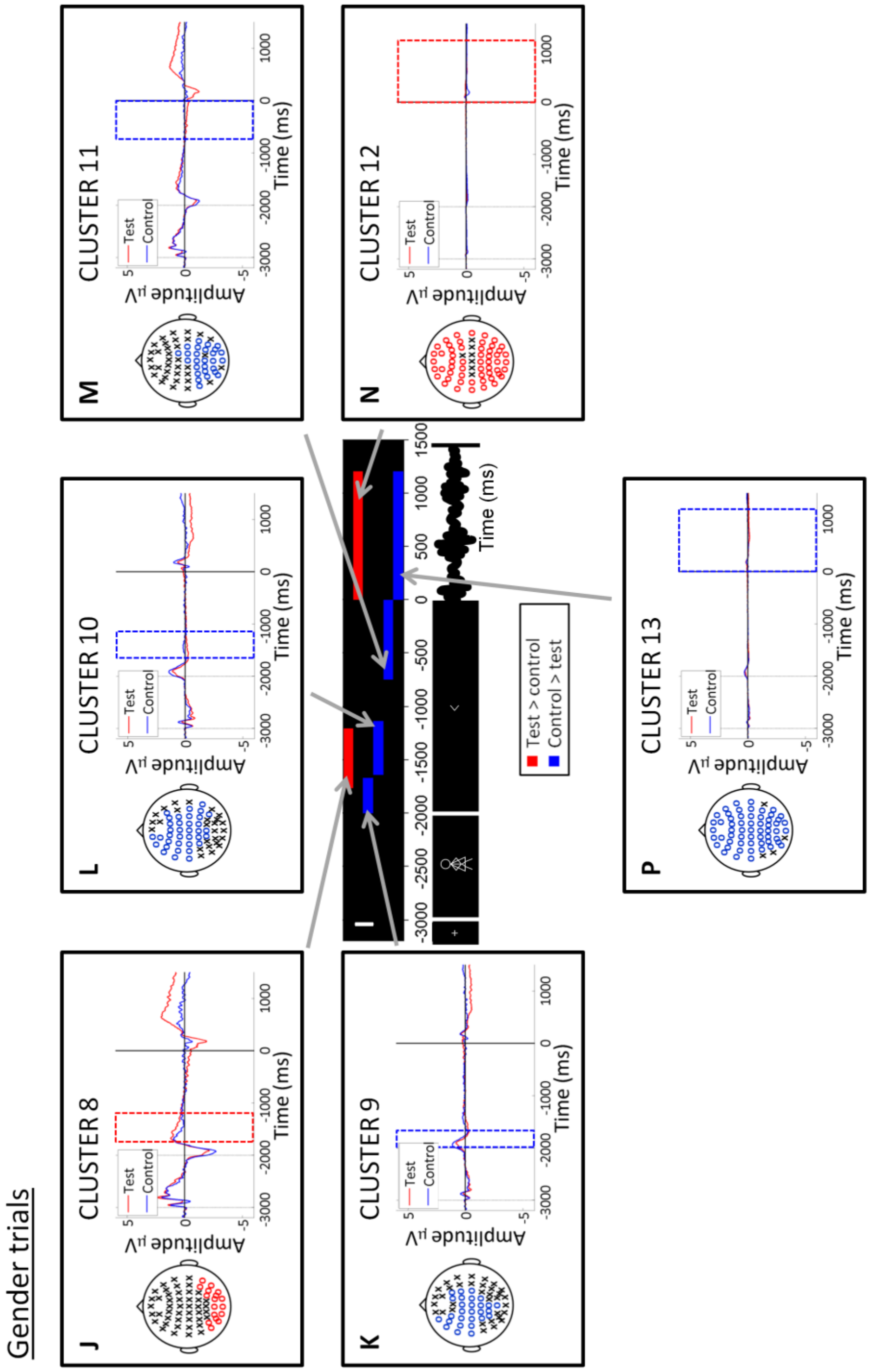


Figure 5.4. (Continued from the previous page)

condition in which a visual cue has no implications for auditory attention and a condition in which the same visual cue directs listeners to prepare to select an upcoming talker on the basis of their gender.

During the Selective Phase, two significant clusters were identified (Figure 5.4I). Cluster 12 (Figure 5.4N) involved 55 electrodes and spanned the entire Selective Phase (0 to 1200 ms). It showed significantly more positive amplitude during the Test Condition than the Control Condition [cluster statistic = 84256, $p < 0.001$]. Cluster 13 (Figure 5.4P; 0 to 1200 ms) was complementary to Cluster 12 in that it spanned the same time window as Cluster 12, but was characterised by the opposite polarity. Cluster 13 involved 58 electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 96637, $p < 0.001$].

5.2.2.2. Type-II analyses: Differences between Location and Gender trials

Differences during the Preparatory Phase

During the Preparatory Phase, four clusters of electrodes were identified that differed significantly in the Test Condition between Location and Gender trials (Figure 5.5A). Cluster 14 (Figure 5.5B; 12 ms to 529 ms) involved 36 mainly posterior electrodes and showed significantly more positive amplitude during Location trials than Gender trials [cluster statistic = 15883, $p = 0.001$]. These values are listed in the first line of Table 5.2. For Cluster 14, the interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [$F(1,19) = 3.03$, $p = 0.10$; Figure 5.6A] and the difference between Location and Gender trials was also present in the Control Condition, $p < 0.001$. When the difference of the differences in Location and Gender trials between the Test and Control conditions was examined in 50-ms sliding windows, the uncorrected p-value was less than 0.05 in only 11 of the 48 50-ms time windows in the cluster (Figure 5.6F). The finding that the pattern of ERPs in this cluster did not differ between the Test and Control Conditions means that it is not possible to rule out the explanation that the cluster arose from differences in the visual cues, rather than from differences in attentional processes triggered by the cues.

Cluster 15 (Figure 5.5C) was complementary to Cluster 14 and was also likely to arise from differences in the visual cues. Cluster 15 (4 to 218 ms) involved 35 central electrodes and showed significantly more negative amplitude in Location trials than Gender trials (cluster statistic = 9328; $p = 0.007$). For this cluster, the interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [$F(1,10) = 0.45$, $p = 0.51$; Figure 5.6B] and the difference between Location and Gender trials was also present in the Control Condition, $p < 0.001$. The uncorrected p-value did not fall below 0.05 during any 50-ms segment over the duration of the cluster (Figure 5.6G).

Table 5.2. (Continued on next page). Summary of results for the Test Condition comparison (Type-II analysis) between Location and Gender trials across Experiments. A tick in the row headed 'Significant in Control Condition?' indicates that the difference in the amplitude of ERPs between Location and Gender trials was significant in the Control Condition across the spatio-temporal points of the cluster (p-values displayed underneath). A tick in the row headed 'Significant Test/Control Interaction?' indicates that an ANOVA with the factors cue type (Location/Gender) and condition (Test/Control) revealed a significant two-way interaction (p-values displayed underneath).

Phase	Properties	Cluster Properties
Preparatory	**Cluster Number**	**14**
	Cluster *p*-value	0.001
	Polarity	Loc > Gen
	Electrode Locations	Posterior
	Onset of cluster (ms)	12
	Duration of cluster (ms)	517
	Significant in Control Condition?	✓ $p < 0.001$
	Significant Test/Control Interaction?	✗ $p = 0.10$
Preparatory	**Cluster Number**	**15**
	Cluster *p*-value	0.007
	Polarity	Gen > Loc
	Electrode Locations	Central
	Onset of cluster (ms)	4
	Duration of cluster (ms)	214
	Significant in Control Condition?	✓ $p < 0.001$
	Significant Test/Control Interaction?	✗ $p = 0.51$
Preparatory	**Cluster Number**	**16**
	Cluster *p*-value	0.020
	Polarity	Gen > Loc
	Electrode Locations	Right Anterior
	Onset of cluster (ms)	168
	Duration of cluster (ms)	355
	Significant in Control Condition?	✓ $p < 0.001$
	Significant Test/Control Interaction?	✗ $p = 0.31$

Table 5.2. (Continued from previous page)

Phase	Properties	Cluster Properties
Preparatory	**Cluster Number**	**17**
	Cluster p-value	0.016
	Polarity	Gen > Loc
	Electrode Locations	Posterior
	Onset of cluster (ms)	835
	Duration of cluster (ms)	251
	Significant in Control Condition?	✗ $p = 0.15$
	Significant Test/Control Interaction?	✗ $p = 0.22$
Selective	**Cluster Number**	**18**
	Cluster p-value	0.030
	Polarity	Loc > Gen
	Electrode Locations	Anterior
	Onset of cluster (ms)	273
	Duration of cluster (ms)	233
	Significant in Control Condition?	✗ $p = 0.85$
	Significant Test/Control Interaction?	✓ p = 0. 037

Cluster 16 (Figure 5.5D; 168 to 523 ms) also showed significantly more negative amplitude in Location Trials than Gender Trials [cluster statistic = 6619, $p = 0.020$] and some of the electrodes overlapped with those identified in Cluster 15. The interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [$F(1,19) = 1.11$, $p = 0.31$; Figure 5.6C] and the difference between Location and Gender trials was also present in the Control Condition, $p < 0.001$. Figure 5.6H shows that the difference between the Test and Control Conditions only reached the $p < 0.05$ (uncorrected) criterion in two 50-ms segments at the end of the cluster. Therefore, it is not possible to rule out the explanation that the cluster arose from differences in the visual cues.

Cluster 17 (Figure 5.5E; 835 to 1086 ms) arose later during the Preparatory Phase. It showed significantly more negative amplitude in Location trials than Gender trials [cluster statistic = 7065, $p = 0.016$] and spanned 20 posterior electrodes. The interaction between cue type (Location/Gender) and condition (Test/Control) was not significant [$F(1,19) = 1.64$, $p = 0.22$; Figure 5.6D], although, unlike Clusters 14–16, the difference between Location and Gender trials was not significant in the Control Condition, $p = 0.15$. Figure 5.6I shows that the difference between the Test and Control Conditions reached the $p < 0.05$ (uncorrected) criterion in nine of the 22 50-ms segments.

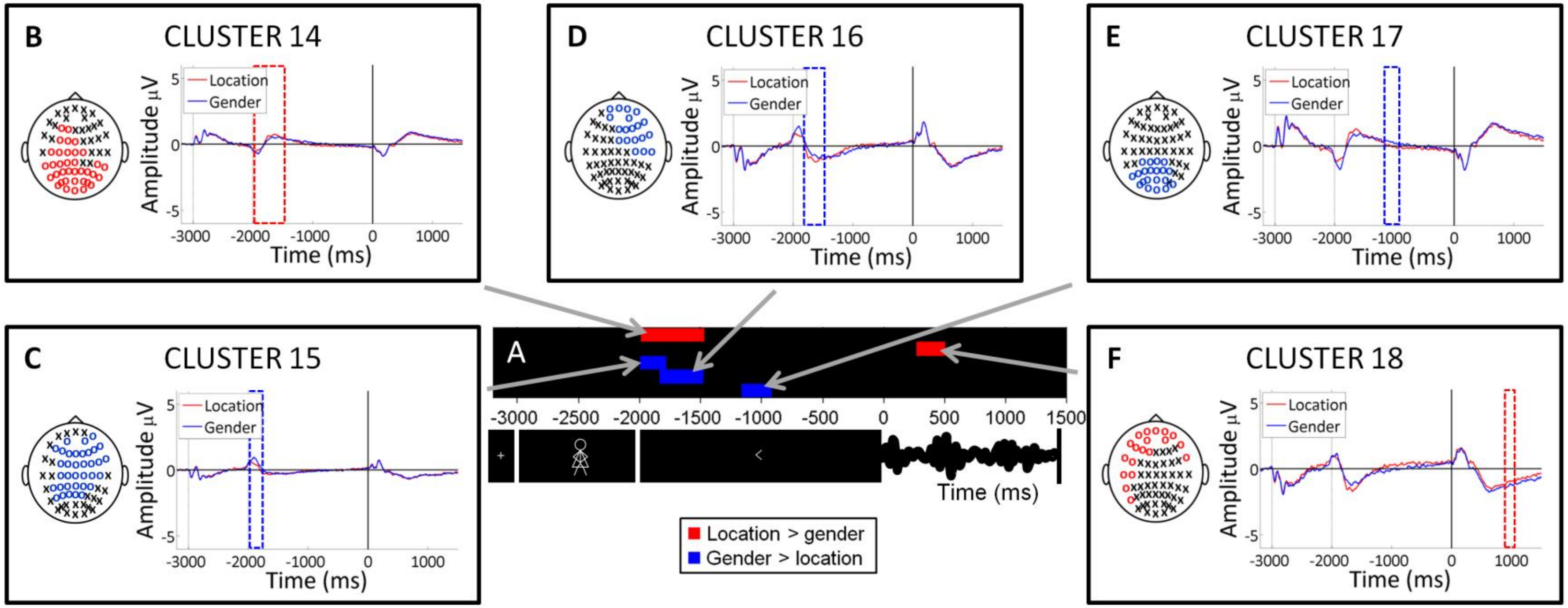


Figure 5.5. Results of the Type-II s patio-temporal Cluster-based Permutation Analysis. This analysis contrasted Location and Gender trials in the Test Condition. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive on location trials than gender trials. For clusters plotted as blue rectangles, the average amplitude was more negative on location trials than gender trials. Further information about each cluster is displayed in (**B**)-(**F**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

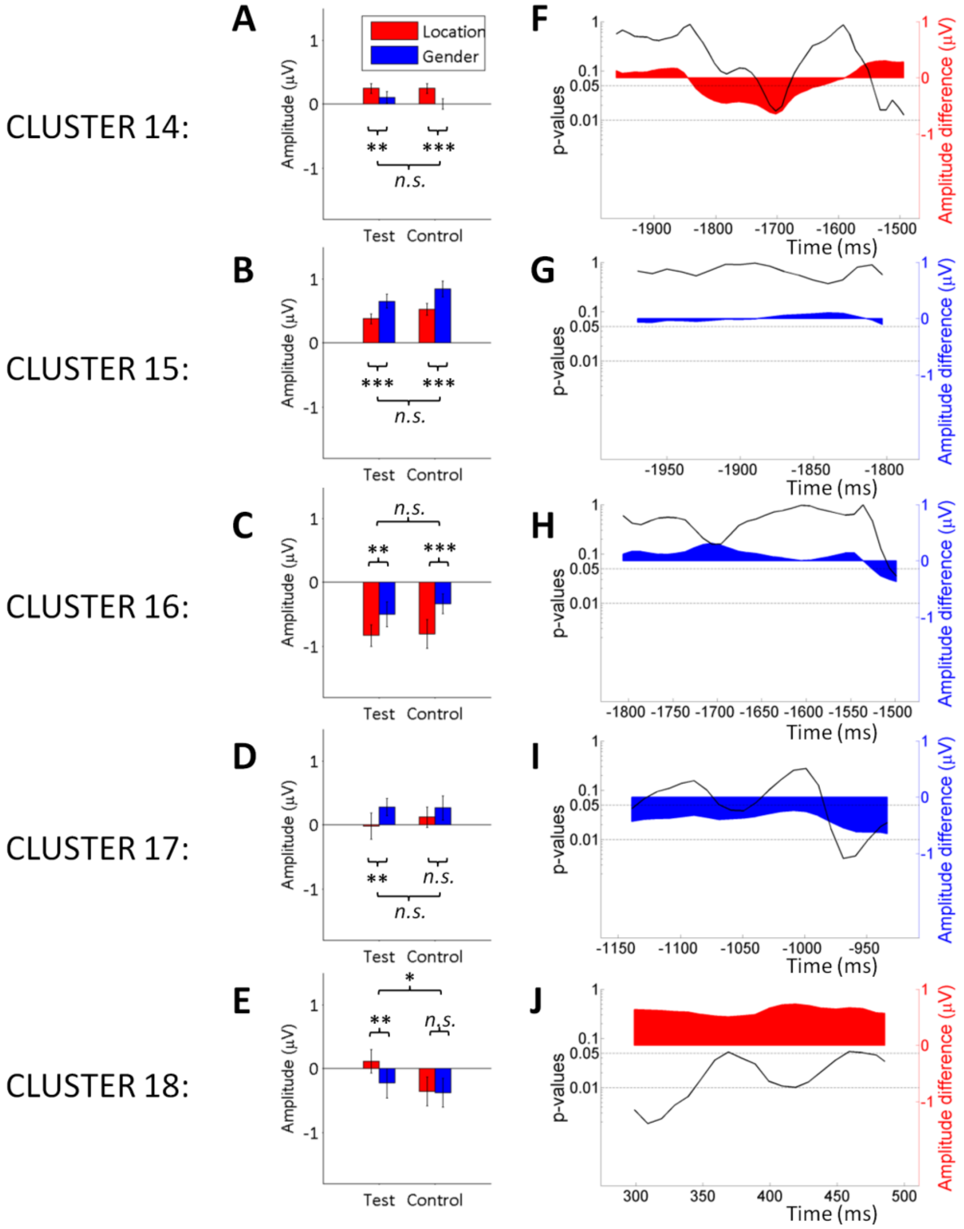


Figure 5.6. Comparison of differences in the amplitude of ERPs between Location and Gender trials in the Test and Control Conditions for each significant Type-II cluster. Graphs (**A**)-(**E**) plot the mean amplitude for Location and Gender trials in the Test and Control Conditions, averaged across participants and space-time points. Error bars show 95% within-subjects confidence intervals. Narrow brackets display the significance level of the comparison between Location and Gender trials in the Test and Control Conditions. Wider brackets display the significance level of the two-way interaction (* $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). Graphs (**F**)-(**J**) display the difference of the differences in Gender and Location trials between the Test and Control conditions in 50-ms time windows repeated every 10 ms within the cluster (right axis) and the uncorrected p-values resulting from a paired-samples t-test comparing the differences (left axis). The mid-point of each time window relative to the onset of acoustic stimuli is displayed on the x-axis.

Differences during the Selective Phase

During the Selective Phase, one cluster of activity was identified that differed significantly between Location and Gender trials (Figure 5.5A). Cluster 18 (Figure 5.5F) lasted from 273 to 506 ms after the onset of the acoustical stimuli. It involved 19 mainly anterior electrode locations and displayed significantly more positive amplitude in Location trials than Gender trials [cluster statistic = 4558, $p = 0.030$]. The interaction between cue type (Location/Gender) and condition (Test/Control) was significant [$F(1,19) = 5.06$, $p = 0.037$; Figure 5.6E] and the difference between Location and Gender trials was not significant in the Control Condition, $p = 0.85$. The uncorrected *p*-value for the difference between Test and Control Conditions was below 0.05 in 50-ms windows throughout most of the cluster (Figure 5.6J; 17 out of 20 50-ms windows) and below 0.001 at the beginning of the cluster. The finding of a significant interaction demonstrates that Cluster 18 arose from differences in the processes for attending selectively to a target talker between Location and Gender trials.

5.2.3. Source reconstruction

Based on the clusters identified in the Cluster-based Permutation Analyses, source reconstruction was performed across three time windows: (1) early during the Preparatory Phase, (2) later during the Preparatory Phase, and (3) during the Selective Phase. The exact time window across which source reconstruction was performed was driven by the exact timing of the clusters from the Spatio-temporal Cluster-based Permutation Analysis of ERPs in each condition.

To ensure that source activity was not cancelled by averaging across the long duration of each time window, paired-samples *t*-tests were conducted on the average amplitude at each voxel that occurred in a 50-ms window that was advanced in 50-ms steps over the duration of each time window. When a Bonferroni correction was applied for multiple comparisons at multiple 50-ms windows and at multiple voxels in the brain, none of the voxels reached the $p < 0.05$ criterion. This implies that there were no differences in the source activity between the Test and Control Conditions or between Location and Gender trials. However, the Bonferroni correction might have been overly stringent, given that activity in consecutive 50-ms windows and at neighbouring voxels is unlikely to be independent. In order to estimate where differences in source activity might possibly occur, a $p < 0.05$ criterion was applied to the uncorrected *p*-values.

Figure 5.7 visualises the comparisons between the Test and Control Conditions, separately for Location and Gender trials. All time windows showed a distributed network of differential activity between the Test and Control Conditions. Early during the Preparatory

Location trials

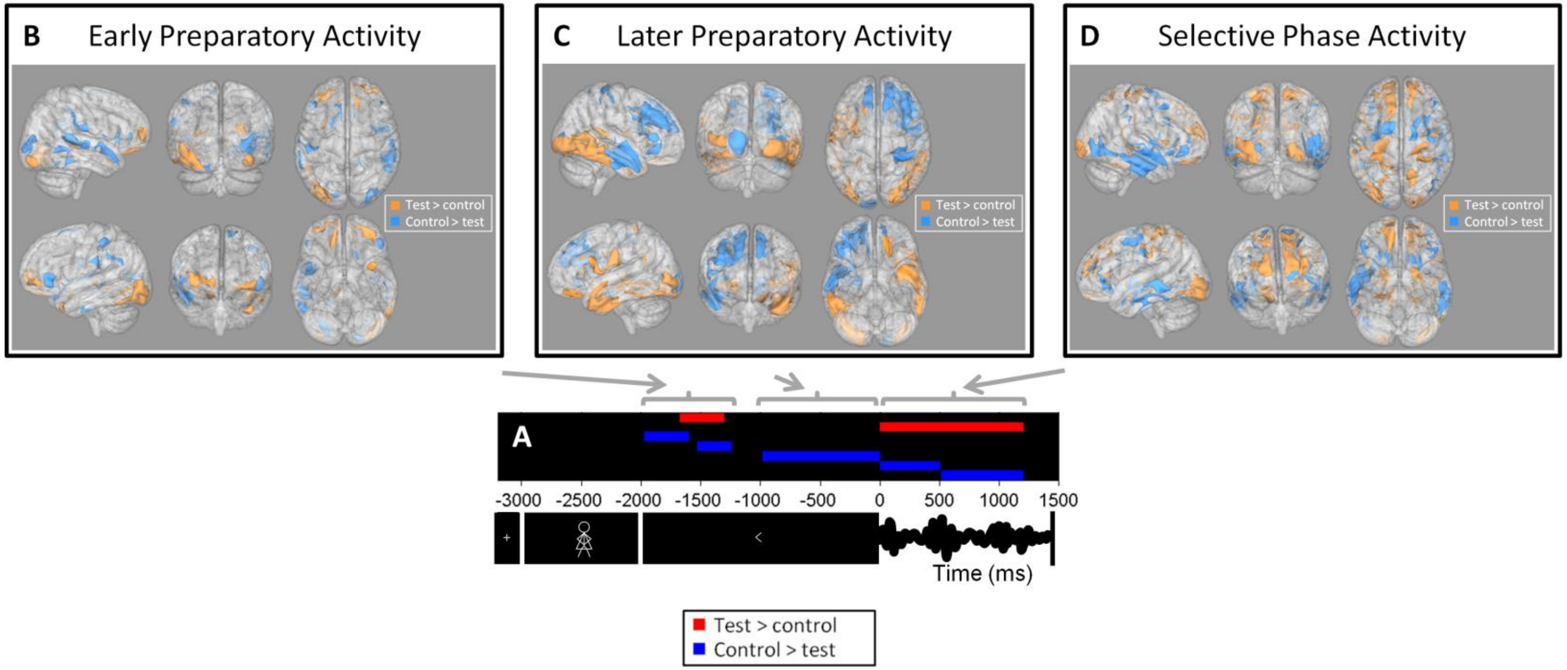


Figure 5.7. (Continued on next page). Results from the source reconstruction comparing the Test and Control Conditions, separately for Location (**A** to **D**) and Gender (**E** to **H**) trials. (**A** and **E**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from the Type-I Spatio-temporal Cluster-based Permutation Analyses for each condition. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant Type-I clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition in the spatio-temporal cluster-based permutation analysis. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Source reconstruction results for three phases of the trial are displayed in (**B** to **D** and **F** to **H**). For each phase, the results of contrasts between the Test and Control Conditions, thresholded at the $p < 0.05$ uncorrected level, are displayed on a 3D view of the MNI 152 average brain. The diagram shows views of the right (top left), left (bottom left), back (top centre), front (bottom centre), top (top right) and bottom (bottom right) of the brain.

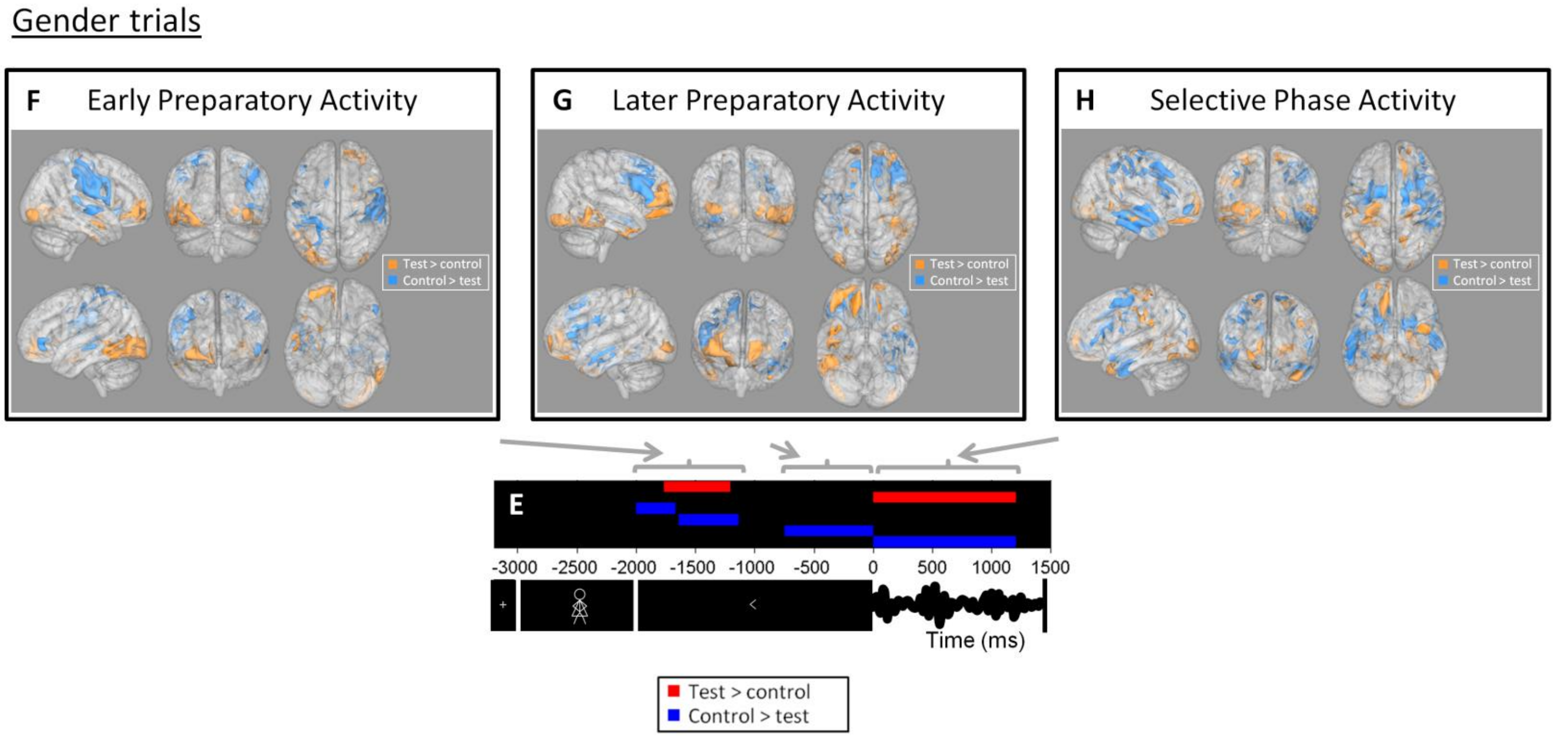


Figure 5.7. (Continued from previous page)

Phase (24 to 762 ms in Location trials and 0 to 860 ms in Gender trials), consistent activity for Location and Gender trials that was greater in the Test Condition than the Control Condition was found in frontal and occipital areas, including parts of medial prefrontal cortex (PFC) and bilateral middle and inferior occipital gyri (Figure 5.7B and F). Consistent activity that was greater in the Control than Test Condition was found in right superior temporal gyrus (STG) and parts of left PFC. Figure 5.8 visualises differential activity between the Location and Gender trials in the Test Condition. Figure 5.8B shows that, early during the Preparatory Phase, there was greater activity in Location than Gender trials in parts of medial PFC and left middle and inferior occipital gyri. In contrast, right STG and the precentral and postcentral gyri showed greater activity in Gender than Location trials.

Later during the Preparatory Phase (1015 to 2000 ms in the Location Condition and 1251 to 2000 ms in the Gender Condition), consistent activity for Location and Gender trials that was greater in the Test Condition than the Control Condition was found in occipital and temporal regions, including bilateral middle and inferior occipital gyri and parts of the right inferior temporal gyrus (Figure 5.7C and G). Consistent activity that was greater in the Control than Test Condition was found in the right middle frontal gyrus. Greater activity in Location than Gender trials was found in right inferior and superior temporal gyri and right inferior and middle occipital gyri (Figure 5.8C). Only small loci of activity showed greater activity in the Gender than Location Condition during this time window.

During the Selective Phase, (0 to 1200 ms for both Location and Gender trials), consistent activity for Location and Gender trials that was greater in the Test Condition than the Control Condition was found in occipital and frontal gyri, including the left middle occipital gyrus and parts of medial PFC (Figure 5.7D and H). Consistent activity that was greater in the Control than Test Condition was found in right inferior, middle and superior temporal gyri. Comparisons between Location and Gender trials (Figure 5.8D) showed greater activity during Location trials in right inferior, middle, and superior temporal gyri and bilateral superior frontal gyrus. In contrast, greater activity during Gender than Location trials was found in parts of left medial frontal gyrus.

5.3. Discussion

The results partially replicate the results of the experiments reported in Chapter 3, which employed a two-talker listening task. The current experiment found significant

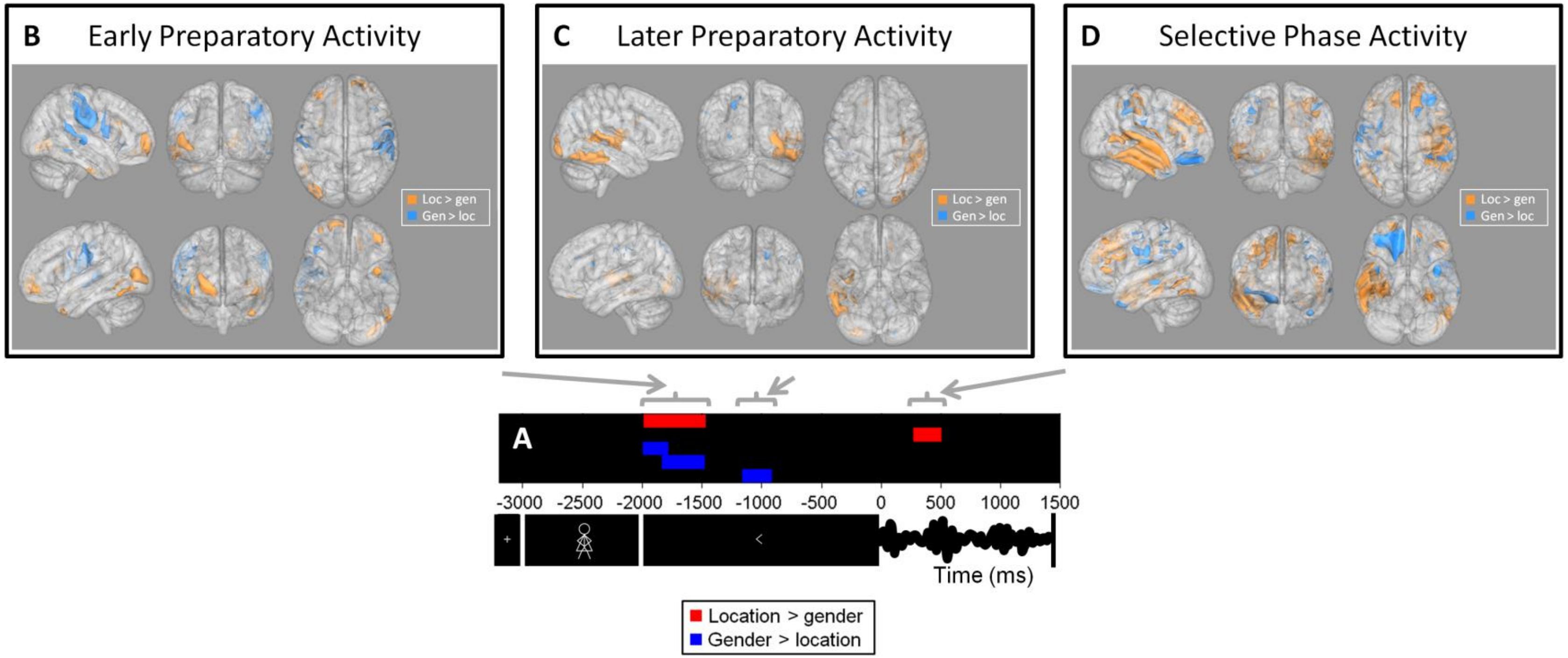


Figure 5.8. Results from the source reconstruction comparing Location and Gender trials in the Test Condition. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from the Type-II Spatio-temporal Cluster-based Permutation Analysis. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant Type-II clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in Location than Gender trials in the Spatio-temporal Cluster-based Permutation Analysis. For clusters plotted as blue rectangles, the average amplitude was more negative in Location than Gender trials. Source reconstruction results for three phases of the trial are displayed in (**B** to **D**). For each phase, the results of contrasts between Location and Gender trials, thresholded at the $p < 0.05$ uncorrected level, are displayed on a 3D view of the MNI 152 average brain. The diagram shows views of the right (top left), left (bottom left), back (top centre), front (bottom centre), top (top right) and bottom (bottom right) of the brain.

differences between the Test and Control Conditions during the Preparatory Phase, separately for Location and Gender trials (Figure 5.4). Since identical visual stimuli were presented in the Test and Control Conditions, this activity can be confidently attributed to attentional preparation for the upcoming task of selecting one of the three talkers based on knowledge of a talker's location or gender. During Location trials, preparatory activity began less than 25 ms after the visual cue was fully revealed. During Gender trials, participants already showed preparatory brain activity by the time the visual cue was fully revealed (significant differences began with 0 ms latency). For *both* Location and Gender trials, additional activity occurred towards the end of the Preparatory Phase, in the 750 ms immediately before the three talkers started speaking.

When comparing trials in which participants attended to Location or Gender, significant differences in EEG activity were observed during the Selective Phase (Figure 5.5), for which differential activity could be attributed to differences in the mechanisms that participants use to pick out a talker based on their location or gender. Differential activity began approximately 250 ms after the onset of acoustical stimuli and lasted approximately 250 ms. The timing of this cluster corresponds to the first half of the acoustical stimuli, when the colour and number key words were spoken.

Accompanying these effects, some additional aspects of activity were likely to result from differences in the visual and acoustical stimuli that were presented in different conditions. For example, differences between the Test and Control Conditions during the Selective Phase are likely to be attributable to differences in the acoustical stimuli, since differential activity was sustained throughout most of the Selective Phase and appeared similar in Location and Gender trials (Figure 5.4). In addition, differences in activity between Location and Gender trials that occurred early (< 600 ms) during the Preparatory Phase were observed in both the Test and Control Conditions (Figure 5.6). Therefore, this result likely reflects differences in physical attributes of the visual cues between Location and Gender trials, such as luminance, structural complexity, or differences in the cognitive processes evoked by animate (human stick figures) and inanimate (chevron) cues (Caramazza & Shelton, 1998; Downing et al., 2006).

5.3.1. Domain-general and cue-specific activity

The finding that ERPs were similar during the Preparatory Phases of Location and Gender trials provides evidence for domain-general preparatory attention. A similar finding was reported in Chapter 3, which reports the results of experiments that used a two-talker task. Together, the results from these experiments provide evidence for domain-general attentional preparation across two different tasks with different acoustical stimuli and which recruited different samples of normally-hearing adults. The finding of domain-general activity is also consistent with the fMRI results reported by Hill and Miller (2010). They found overlapping activity in a left-dominant fronto-parietal network in response to a visual cue for location or F0, before three talkers started speaking.

The comparison between Location and Gender trials aimed to reveal whether there was additional cue-specific processing, as reported by Hill and Miller (2010) and Lee *et al.* (2013). The early-latency clusters (< 600 ms) that differed between Location and Gender trials during the Preparatory Phase were likely to result from differences in physical attributes of the visual cues between Location and Gender trials (Figure 5.6A–C). However, it is possible that the later-latency cluster (which began later than 800 ms after the visual cue was revealed), reflects cue-specific attentional preparation for a talker based on location or gender information. For this cluster (Cluster 17), a significant difference between Location and Gender trials was not present in the Control Condition, in which the same cues were presented when they had no implications for auditory attention (Figure 5.6D). However, this result should be interpreted with caution given that a similar *pattern* of amplitudes (albeit with a slightly smaller difference in amplitude between Location and Gender trials) emerged in the Control Condition.

Differences between Location and Gender trials during the Selective Phase revealed differential activity that could be attributed to differences in the mechanisms that participants use to pick out a talker based on their location or gender. A similar finding was reported in Chapter 3 for a two-talker listening task. The finding of cue-specific activity during the Selective Phase is consistent with the results of Hill and Miller (2010) and Lee *et al.* (2013), who both found significant differences in brain activity when participants selectively attended to a talker depending on whether participants received information about the talker's spatial location or their F0.

5.3.2. Localisation of source activity

The neural generators of EEG activity were not able to be specified with statistical precision and, as a result, the analyses estimated where differences in source activity might occur. Therefore, the results may reflect high activity in only a small number of participants, rather than effects that were consistent across all participants, and should be interpreted with this limitation in mind. A replication using more sophisticated analyses would be desirable before firm conclusions are drawn.

5.3.2.1. Similar activity in Location and Gender trials

There were several areas in which greater activity for the Test than Control Condition was observed in *both* Location and Gender trials. Across all time windows, greater activity was observed for the Test than Control Condition in occipital gyri and the inferior temporal gurus (Figure 5.7). These findings are consistent with the idea that participants had to interpret the visual cues in the Test Condition in order to use them to prepare auditory attention. The inferior temporal gurus has been implicated in visual object recognition (e.g. Denys et al., 2004; Gross, Rocha-Miranda, & Bender, 1972; Gross, 1992) so may have been involved in interpreting the cues, which were symbolic. During the Selective Phase and early during the Preparatory Phase, the medial PFC also showed greater activity in the Test than Control Condition (Figure 5.7). This result is consistent with the well-established role of the PFC in top-down attention (e.g. Gregoriou, Rossi, Ungerleider, & Desimone, 2014; Lebedev, Messinger, Kralik, & Wise, 2004; Salmi et al., 2007).

There were also similarities in the regions that showed greater activity in the Control than Test Condition in Location and Gender trials (Figure 5.7), although these results are difficult to interpret. One possible explanation is that they reflect greater activity related to maintenance of the visual cue in Control Condition, since in the Control Condition, participants task is to identify the visual cue presented on each trial, but in the Test Condition, participants need to use the cue but do not respond to it.

The finding of similar source estimates in Location and Gender trials suggests that the sources identified were more robust than the statistics otherwise suggest. In addition, overlapping activity in Location and Gender trials is consistent with Hill and Miller's (2010) fMRI results. During the Preparatory and Selective Phases of their task, they found BOLD activity in overlapping networks when participants received cues for location (left/right) and F0 (high/low). Therefore, both experiments provide evidence for domain-general attentional processing when participants were cued to different attributes of a talker (location and gender in the current experiment and location and F0 in Hill and Miller's experiment).

5.3.2.2. Differential activity between Location and Gender trials

Accompanying overlapping activity, the results also revealed some areas in which activity was likely to differ between Location and Gender trials. During the Preparatory Phase, there was greater activity for Location than Gender trials in visual cortices (Figure 5.8B–C). This finding is likely to reflect differences in physical aspects of the visual cues between Location and Gender trials. For example, differences in luminance, complexity, or differences in the cognitive processes evoked by the representation of an inanimate object (a chevron) compared with a human being. The finding of activity in visual areas during the

Preparatory Phase is consistent with the results of the Spatio-temporal Cluster-based Permutation Analysis, which identified differences in ERPs between Location and Gender trials in both the Test and Control Conditions that were attributed to visual processing of the cues.

In addition, early during the Preparatory Phase, the medial PFC showed greater activity in Location than Gender trials (Figure 5.8B). The finding of differential activity in prefrontal cortex is consistent with the finding that spatial and non-spatial *visual* information can be represented in different areas of the primate prefrontal cortex (e.g. Wilson, O'Scaoaidhe, & Goldman-Rakic, 1993).

5.3.2.3. Correspondence with previous multi-talker listening experiments

Overall, there were both similarities and differences between the regions identified in the current experiment and regions identified in the previous experiments of Hill and Miller (2010) and Lee *et al.* (2013). One possible explanation for discrepancies is that differences in the stimuli presented in the current experiment compared to the experiments of Hill and Miller and Lee *et al.* might contribute to differences in the patterns of neuronal activity observed. First, Hill and Miller and Lee *et al.* cued attention to F0, rather than gender. Although F0 is one factor that contributes to differentiation of talkers by gender, it is possible that *either* subtle differences in the ways in which participants utilise cues for F0 and gender *or* differences in the visual cues used to cue attention to these attributes (upwards- and downwards-pointing chevrons in the previous experiments but stick figures in the current experiment) might underlie different estimates of source activity. Second, the previous experiments directed attention to sounds presented in virtual space through headphones, which might evoke a different pattern of neuronal activation than attending to speech presented in different locations in the sound field.

Overall, it is important to consider the limitation that sources in the current experiment were not able to be identified with statistical precision. This factor might also contribute to differences in estimates of source location between the current experiment and the experiments of Hill and Miller and Lee *et al.*

5.3.3. Conclusions

In summary, this experiment provides evidence for domain-general and cue-specific EEG activity during three-talker listening. Preparatory attention for a talker's location or gender began early (< 25 ms) after the visual cue was revealed and was sustained throughout the Preparatory Phase. During the Selective Phase, there was evidence for cue-specific EEG activity that depended on whether the listener attended selectively to a talker on the basis of knowledge of their location or gender.

Overall, this experiment provides two main contributions. First, it improves understanding of the time course of brain activity in normally-hearing adults during multi-talker listening. Second, it identifies and tests a technique for measuring attentional brain activity during multi-talker listening that has the potential to be applied in future experiments in order to identify atypical attentional processing. The experiments reported in the next chapter exploited this potential to seek differences in preparatory attention between normally-hearing and hearing-impaired listeners.

Chapter 6

Auditory Attention in Children with Hearing Loss

Converging evidence from accuracy scores and self-report suggests that multi-talker listening is particularly challenging for listeners with sensorineural hearing loss (e.g. Dubno et al., 1984; Helfer & Freyman, 2008). However, the central consequences of hearing loss, including how hearing-impaired listeners direct attention to speech in noisy environments, are not fully understood. These experiments aimed to investigate one possible central consequence of hearing loss: difficulties with preparatory attention. The experiments reported in this chapter investigated (1) how the duration of preparation time affects the accuracy and latency of speech intelligibility for normally-hearing and hearing-impaired children, and (2) the event-related potentials (ERPs) evoked by these groups of participants during three-talker listening.

Investigating speech intelligibility and ERPs in *children* with hearing loss is particularly interesting because, unlike older adults who experience peripheral hearing loss as a consequence of normal ageing, differences in peripheral acoustical processing are not confounded with a general cognitive decline with older age that is separate from hearing loss itself. Furthermore, the children with hearing loss who took part in these experiments had early-onset or congenital hearing loss. As a result, the input from the periphery to the brain would have been distorted for most of, if not their entire, lives. Therefore, they may not have experience of using the cues that normally-hearing listeners deploy to segregate simultaneous talkers, such as cues for location or gender.

Two of the experiments report results for normally-hearing children who belonged to a similar age range as the hearing-impaired children, with the aim of providing a comparison group for interpreting the results from hearing-impaired children. Dhamani, Leung, Carlile, and Sharma (2013) showed that normally-hearing children, like adults, can use advance cueing to improve the accuracy of speech intelligibility in noisy environments. They asked children aged 10–15 years to identify a target syllable in a background of two-talker babble. On each trial, an auditory priming cue was presented in advance of the target array and indicated the onset time of the target syllable. Accuracy of reporting the target syllable was better when the cue validly predicted onset time, compared to when the target was presented earlier or later than expected. This result shows that children aged 10–15 years are able to direct their attention to a talker based on a cue that indicates the onset time of the talker. In addition, the results reported in Chapter 3 showed moderately high accuracy of speech intelligibility for normally-hearing children aged 7–13 years in a two-talker listening task that was similar to the three-talker task employed in the current experiments. Therefore, children were expected to display similar patterns of results as those observed in adults in the previous chapters.

6.1. Speech intelligibility during three-talker listening

The first two experiments measured the accuracy and latency of speech intelligibility during three-talker listening in normally-hearing (Experiment 1) and hearing-impaired (Experiment 2) children. The experiments employed a three-talker listening task that was similar to the task used in Experiment 2 of Chapter 4. Before the three talkers began to speak, a visual cue indicated either the location (left/right) or the gender (male/female) of the target talker. Participants' task was to report key words spoken by the target talker. The interval between the time that the cue for location or gender was revealed and the time that the talkers started speaking (i.e. the cue-target interval) was varied between 0 and 2000 ms. Experiments 1 and 2 aimed to investigate whether the accuracy and latency of speech intelligibility was improved when participants had more time to prepare for the location or for the gender of an upcoming talker.

One previous experiment found that the accuracy of speech intelligibility did not improve as much for hearing-impaired adults as for normally-hearing adults when they received an advance cue for location (Best et al., 2009). This result is consistent with the idea that hearing-impaired listeners do not utilise preparatory attention in the same way as normally-hearing listeners. One possibility is that hearing-impaired listeners do not deploy preparatory attention to the same *extent* as normally-hearing listeners. In this case, hearing-impaired children would be expected to show a smaller improvement in the accuracy of speech intelligibility between cue-target intervals of 0 and 2000 ms than normally-hearing children. An alternative possibility is that hearing-impaired listeners require *more time* to prepare effectively for an upcoming talker than do normally-hearing listeners. In this case, hearing-impaired children would be expected to show an improvement in speech intelligibility that is equivalent to normally-hearing children at longer, but not at shorter, cue-target intervals.

6.2. ERPs during three-talker listening

Experiments 3 and 4 measured ERPs during a three-talker listening task in which the duration of preparation time was fixed at 2000 ms. It was expected that normally-hearing children (Experiment 3) would display a similar pattern of ERPs as the normally-hearing adults reported in Chapter 5. This hypothesis led from the results reported in Chapter 3, in which normally-hearing children showed similar patterns of ERPs as adults in a two-talker listening task.

Experiment 4 measured ERPs evoked during three-talker listening in children with moderate bilateral hearing loss of cochlear origin. Experiment 4 was expected to reveal ERPs that were different to those measured from normally-hearing children. Atypical ERPs during the acoustical stimuli (i.e. the Selective Phase) were expected to reflect impaired peripheral transduction in hearing-impaired listeners. Atypical ERPs during the Preparatory Phase (i.e. between the reveal of a visual cue and the onset of the talkers) were expected to reflect atypical attentional preparation for an upcoming talker based on cues for location or gender. Importantly, the design of the experiment ensured that differences during the Preparatory Phase could not be explained by differences in transduction at the auditory periphery because no acoustical stimuli were presented during the Preparatory Phase.

6.3. Experiment 1

6.3.1. Methods

6.3.1.1. Participants

Participants were 20 children (9 male), aged 7–16 years (mean [M] = 11.0, standard deviation [SD] = 2.1). All participants were declared by their parents to be native English speakers with no history of hearing problems. They had 5-frequency average pure-tone hearing levels of 20 dB HL or better, tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). Fifteen of the children had previously taken part in Experiment 3

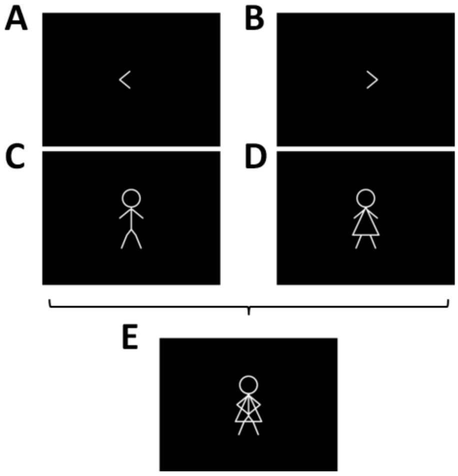


Figure 6.1. (**A**)-(**D**) Visual cues. (**E**) Visual composite stimulus, which is a combination of the four visual cues overlaid.

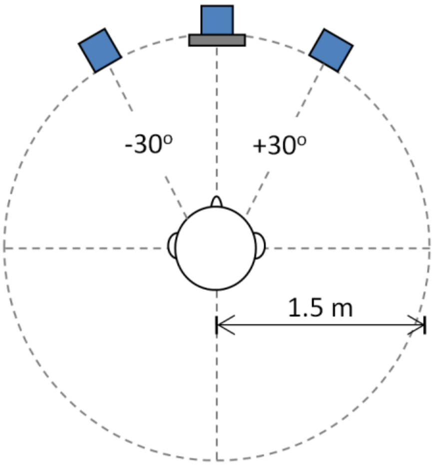


Figure 6.2. Layout of loudspeakers (blue squares) and visual display unit (grey rectangle) relative to a participant's head.

of Chapter 3. The study was approved by the Research Ethics Committee of the Department of Psychology of the University of York.

6.3.1.2. Apparatus and stimuli

The apparatus and stimuli were the same as Experiment 2 of Chapter 4 (the visual stimuli are illustrated in Figure 6.1), except that the left and right loudspeakers were located at ± 30° azimuth (Figure 6.2). The average presentation level of concurrent triplets of test sentences was set to 63 dB(A) (range 61.6—66.2 dB) measured with a B&K (Brüel & Kjær, Nærum, Denmark) Sound Level Meter (Type 2260 Investigator) and 0.5-inch Free-field Microphone (Type 4189) placed in the centre of the arc at the height of the loudspeakers with the participant absent.

6.3.1.3. Procedure

Figure 6.3 shows the trial structure, which was similar to Experiment 2 in Chapter 4. The only difference was that, due to time constraints, only three cue-target intervals were presented: 0, 1000, and 2000 ms. Participants completed between 192 and 288 trials (depending on their level of fatigue). Each participant completed an equal number of trials for each of the three cue durations and, for each duration, an equal number of trials for each of the different visual cues. Participants received a short break every 16 trials and a longer break every 48 trials.

Participants completed two sets of familiarisation trials before the main task, similar to Experiment 2 reported in Chapter 4. In the first set, 12 trials were presented, in which *either* the male or female talker was presented on each trial from either the left or right

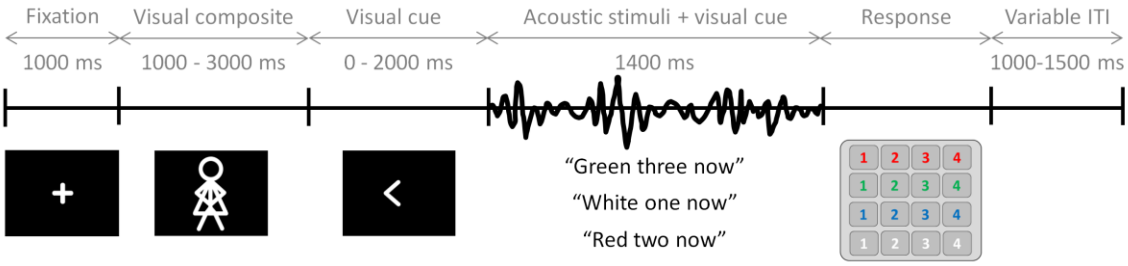


Figure 6.3. Trial structure of Experiments 1 and 2, with an example trial below.

loudspeaker. In the second set, 4 trials were presented, in which all three voices were presented on each trial. The second set of trials was identical to the main task.

6.3.1.4. Analyses

Accuracy and RTs

Trials were separated into attend-location (average left/right cues) and attend-gender (average male/female cues) groups, separately for each of the three cue-target interval conditions. For each trial, three categories of response were recorded: (1) correct identification of both the colour and number (i.e. the "Colour-number combination") spoken by the target talker; (2) correct identification of the colour irrespective of whether the number was reported correctly ("Colour-only"); (3) correct identification of the number irrespective of whether the colour was reported correctly ("Number-only"). In addition, reaction times (RTs), measured from the onset of the acoustical stimuli, were calculated on trials in which participants correctly identified the Colour-number combination. RTs beyond two standard deviations from the mean for each participant were excluded from the analysis.

Errors

When participants did not correctly identify the Colour-number combination, responses were categorised into one of four different types of error. The reported Colour-number combination could be: (1) spoken by the opposite-gender talker that was presented from the contralateral location ("opposite-gender" error), (2) spoken by the "child" talker that was presented from the central location ("child" error), (3) a mixture of words spoken by the target and a non-target talker *or* a mixture of words spoken by the two non-target talkers ("mix" error), or (4) not be spoken by any mixture of the talkers on that trial ("absent" error).

The percentages of the four types of error were assessed in relation to the percentages expected if participants guessed randomly with a uniform distribution. The expected percentages were: 6.7% "opposite-gender" error, 6.7% "child" error, 40.0% "mix" error, and 46.7% "absent" error.

6.3.2. Results

6.3.2.1. Colour-number accuracy

Figure 6.4A illustrates the percentages of trials on which participants correctly reported the Colour-number combination that was spoken by the target talker. A 3 x 2 repeated-measures ANOVA with the factors cue-target interval (0/1000/2000 ms) and cue type (location/gender) showed a significant main effect of cue-target interval, $F(2, 38) = 16.23$, $p < 0.001$, $\eta_p^2 = 0.46$. Contrasts showed that the 1000-ms [$F(1, 19) = 18.92$, $p < 0.001$, $\eta_p^2 = 0.50$] and 2000-ms [$F(1, 19) = 24.70$, $p < 0.001$, $\eta_p^2 = 0.57$] intervals both led to significantly higher Colour-number accuracy than the 0-ms interval. A Bonferroni-corrected post-hoc test showed no significant difference between the 1000-ms and 2000-ms intervals. There was no significant main effect of cue type [$F(1, 19) = 0.03$, $p = 0.86$] and no significant interaction between cue-target interval and cue type [$F(2, 38) = 0.76$, $p = 0.48$].

6.3.2.2. Colour-only accuracy

The pattern for Colour-only accuracy (Figure 6.4B) was similar to the pattern for Colour-number accuracy. There was a significant main effect of cue-target interval, $F(2, 38) = 10.20$, $p < 0.001$, $\eta_p^2 = 0.35$. Similarly, the 1000-ms [$F(1, 19) = 5.40$, $p = 0.031$, $\eta_p^2 = 0.22$] and 2000-ms [$F(1, 19) = 19.43$, $p < 0.001$, $\eta_p^2 = 0.51$] intervals led to significantly higher colour accuracy than the 0-ms interval. There was no significant main effect of cue type [$F(1, 19) = 0.03$, $p = 0.87$] and no significant interaction between cue-target interval and cue type [$F(2, 38) = 0.61$, $p = 0.55$].

6.3.2.3. Number-only accuracy

Figure 6.4C illustrates the percentages of trials on which participants correctly reported the number that was spoken by the target talker. There was a significant main effect of cue-target interval, $(2, 38) = 4.96$, $p = 0.012$, $\eta_p^2 = 0.21$. Only the 2000-ms interval led to significantly higher Number-only accuracy than the 0-ms interval [$F(1, 19) = 9.12$, $p = 0.007$, $\eta_p^2 = 0.32$]. There was no significant main effect of cue type [$F(1, 19) = 0.02$, $p = 0.88$] and no significant interaction between cue-target interval and cue type [$F(2, 38) = 1.43$, $p = 0.25$].

6.3.2.4. RTs

Figure 6.4D illustrates the RT results. There was a significant main effect of cue-target interval, $F(2, 38) = 476.94$, $p < 0.001$, $\eta_p^2 = 0.96$. Contrasts showed that the 1000-ms [$F(1, 19) = 773.90$, $p < 0.001$, $\eta_p^2 = 0.98$] and 2000-ms [$F(1, 19) = 746.24$, $p < 0.001$, $\eta_p^2 = 0.98$] intervals both produced significantly shorter RTs than the 0-ms cue-target interval. Bonferroni-corrected post-hoc tests also showed significantly longer RTs at the 2000-ms than the 1000-ms interval ($p < 0.001$).

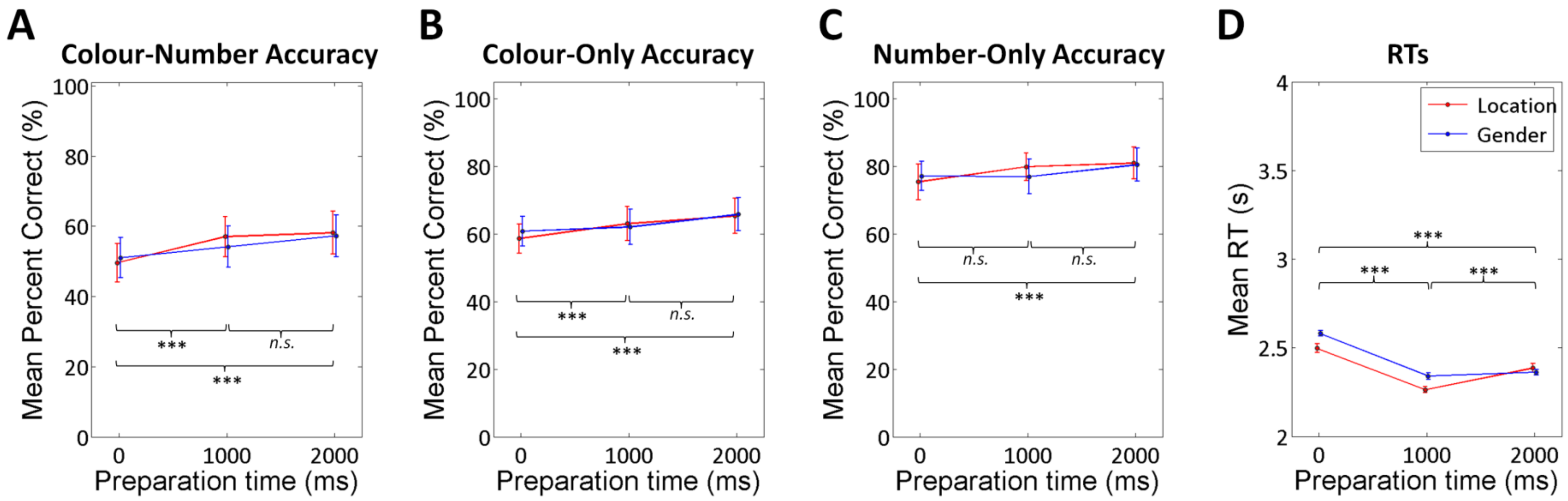


Figure 6.4. Accuracy and reaction time (RT) results for Experiment 1. (**A**) Mean percentage of trials in which participants correctly identified the Colour-number combination spoken by the target talker. (**B**) Mean percentage of trials in which participants correctly identified the colour (irrespective of the number). (**C**) Mean percentage of trials in which participants correctly identified the number (irrespective of the colour). **(D**) Mean RTs on trials in which participants correctly identified the colour-number combination. Error bars show within-subjects 95% confidence intervals. Brackets indicate the significance level of contrasts comparing 0-ms and 1000-ms and 0-ms and 2000-ms and post-hoc tests comparing 1000-ms and 2000-ms (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$).

RTs were significantly shorter in the attend-location condition (M = 2.38 s, SD = 0.17) than the attend-gender condition (M = 2.43 s, SD = 0.15), $F(1, 19) = 18.50$, $p < 0.001$, $\eta_p^2 = 0.49$. There was also a significant two-way interaction between cue-target interval and cue type, $F(2, 38) = 20.16$, $p < 0.001$, $\eta_p^2 = 0.52$. Bonferroni-corrected post-hoc tests showed significant differences between all cue-target intervals, separately for attend-location and attend-gender conditions ($p \leq 0.001$), except for the difference between the 1000-ms and 2000-ms intervals in the attend-gender condition ($p = 0.20$)

6.3.2.5. Errors

The largest percentages of errors were "opposite-gender" errors (M = 37.2%, SD = 11.5) and "child" errors (M = 33.4%, SD = 14.5). The percentages of "mix" errors (M = 16.1%, SD = 6.5) and "absent" errors (M = 13.3%, SD = 7.7) were smaller. The percentages of "opposite-gender" [$t(19) = 11.21$, $p < 0.001$] and "child" [$t(19) = 8.30$, $p < 0.001$] errors were significantly greater than their expected values, whereas the percentages of "mix" [$t(19) = 17.03$, $p < 0.001$] and "absent" [$t(19) = 21.87$, $p < 0.001$] errors were significantly smaller than their expected values.

A 4 x 3 x 4 repeated-measures ANOVA investigated whether the percentages of different error types (4 levels: "opposite-gender", "child", "mix", and "absent" errors) differed significantly between the different cue-target intervals (3 levels: 0, 1000, and 2000 ms) or between cue types (4 levels: left, right, male, and female). There was a significant main effect of error type, $F(1.7, 57) = 21.18$, $p < 0.001$, $\eta_p^2 = 0.53$. Bonferroni-corrected post-hoc tests showed that the percentage of "opposite-gender" errors did not differ significantly from the percentage of "child" errors ($p > 0.99$), the proportion of "mix" errors did not differ

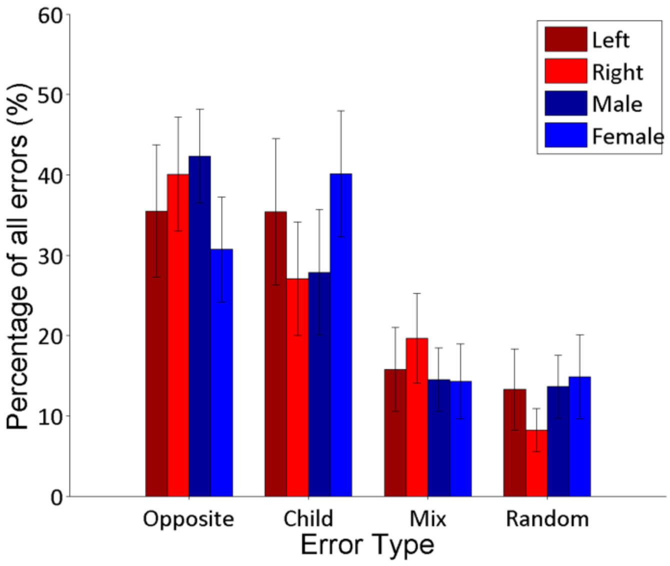


Figure 6.5. Experiment 1: Comparison of the percentages of incorrect responses that were categorised into each of the four error-types ("opposite-gender", "child", "mix", or "random" errors) for each of the four different cue types (left/right/male/female).

significantly from the percentage of "absent" errors ($p = 0.83$), but there were significant differences between the percentages of all other error type combinations ($p \leq 0.005$).

There was no significant difference in the proportion of errors made for different cue-target intervals [error type * cue-target interval interaction: $F(3.8, 71.9) = 0.34$, $p = 0.84$], but there was a significant difference in the proportion of errors made across the four different cue types [Figure 6.5; error type * cue type interaction: $F(5.3, 99.9) = 4.39$, $p = 0.001$, $\eta_p^2 = 0.19$]. However, when the cue-type variable was collapsed into attend-location (average left/right) and attend-gender (average male/female) trials, there were no significant differences in the proportion of errors made between attend-location and attend-gender trials overall [error type * attend location/gender interaction: $F(3, 57) = 0.78$, $p = 0.51$].

6.3.3. Discussion

Colour-number and Colour-only accuracy were higher when participants received cue-target intervals of 1000 or 2000 ms compared to 0 ms (i.e. when the cue was revealed at the same time as the acoustical stimuli began). Number-only accuracy was also higher at the 2000-ms than the 0-ms cue-target interval. This result shows that normally-hearing children achieve better speech intelligibility when they have time to prepare for the location or gender of an upcoming target talker than when they have no time to prepare before the talkers start speaking.

However, a possible alternative explanation for better accuracy for intervals of 1000 and 2000 ms than 0 ms is that the 0-ms interval produced a significant detriment to speech intelligibility (rather than an improvement in intelligibility for intervals greater than 0 ms). A possible argument is that presenting a simultaneous visual stimulus might distract attention from the acoustical stimuli (which participants had to report) in the 0-ms condition. However, this alternative explanation is unlikely to account for the current results. First, during the same task, adults showed a *progressive* improvement in the latency of speech intelligibility as the duration of preparation time increased up to 2000 ms (reported in Experiment 2 of Chapter 4). Second, Experiment 3 of Chapter 3 demonstrates that children of a similar age to those in the current experiment showed significant EEG activity during the 1000-ms Preparatory Phase of a two-talker listening task. Those findings are inconsistent with the explanation that distraction in the 0-ms condition underlies differences in speech intelligibility between the 0-ms and the 1000-ms and 2000-ms intervals. Rather, those results support the explanation that the children in Experiment 1 deployed preparatory attention in the interval between the reveal of the visual cue and the onset of the talkers.

The patterns in the RT data were somewhat similar to the accuracy results. RTs were shorter at the 1000-ms and 2000-ms intervals compared to the 0-ms interval. However, there was also a significant difference in the RTs between 1000 and 2000 ms in the direction opposite to the prediction—RTs were longer for the 2000-ms interval than the 1000-ms interval. Since accuracy improved between 1000 and 2000 ms intervals, it is possible that this result reflects a speed-accuracy trade-off.

Accuracy and errors were not affected overall by whether participants received a cue for location or gender. However, RTs were significantly faster on attend-location than attend-gender trials. The RT data also showed a significant interaction between attention to location or gender and the cue-target interval. The significant interaction appears to be driven by shorter RTs for attend-location than attend-gender trials at 0 and 1000 ms but similar RTs at 2000 ms. This pattern of results demonstrates that children required more preparation time when they received the gender cue to produce RTs equivalent to the location cue condition.

6.4. Experiment 2

Experiment 2 aimed to investigate whether children with moderate bilateral hearing loss showed the same pattern of accuracy and RT results as normally-hearing children when the duration of preparation time varied between 0 and 2000 ms. Experiment 2 used a similar task as Experiment 1. The only difference was the presentation level of the acoustical stimuli, which was increased from 63 to 76 dB(A) SPL in order to compensate, in part, for the elevated pure-tone thresholds of participants with moderate bilateral hearing loss. Hearing-impaired children were expected to show a different pattern of results to the normally-hearing children whose results are reported in Experiment 1.

In addition, the words that hearing-impaired children reported on incorrect trials were used to make inferences about the aspects of three-talker listening that they struggle with. On each trial, participants had to report two key words that were spoken by the target talker. If hearing-impaired children showed poor speech intelligibility because they were not able to segregate words spoken by the target talker from those of competing talkers (i.e. difficulties in 'object formation'), it was expected that they should produce a higher percentage of errors that consisted of a mixture of words spoken by the target and/or competing talkers on each trial than normally-hearing children. However, if poor speech intelligibility was a result of difficulties in selecting the correct talker to attend (i.e. difficulties in 'object selection'), independent of problems of object formation, then hearing-impaired children would be expected to show a higher percentage of errors spoken by one of the competing talkers. Finally, if poor speech intelligibility was a result of higher energetic masking, then hearing-impaired children would be expected to be more likely to report words from the array of possible words that were not spoken by any of the talkers on each trial.

6.4.1. Methods

6.4.1.1. Participants

Participants were 9 children (1 male), aged 9–16 years (M = 12.1, SD = 2.2). Eight children had moderate hearing loss and one child had mild hearing loss. The children were identified as having sensorineural hearing loss by the audiologist at the hospital at which they receive care, although a more detailed etiology is unknown. They had 5-frequency average pure-tone hearing levels between 30 and 61 dB HL (M = 50 dB HL, SD = 9.6), tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). The difference in the 5-frequency averages recorded from the left and right ears was less than 8 dB for each participant. Participants were declared by their parents to be native English speakers. Out of the nine children, one had an additional visual impairment in her left eye. All participants had taken part in Experiment 4 (reported below) before taking part in this experiment. Participants completed the experiment without using their hearing aids. The study was approved by the Research Ethics Committee of the Department of Psychology of the University of York.

6.4.1.2. Apparatus, stimuli, procedure, and analyses

The apparatus, stimuli, procedure, and analyses were the same as Experiment 1, except for the presentation level of the acoustical stimuli. The average presentation level of concurrent triplets of test sentences was set to 76 dB(A) (range 72.4—77.9 dB) measured with a B&K (Brüel & Kjær, Nærum, Denmark) Sound Level Meter (Type 2260 Investigator) and 0.5-inch Free-field Microphone (Type 4189) placed in the centre of the arc at the height of the loudspeakers with the participant absent. All participants completed 288 trials during the main task, after the familiarisation trials had been administered.

Additional analyses directly compared the results from the current set of participants with the results gathered from normally-hearing children in Experiment 1.

6.4.2. Results

6.4.2.1. Results from hearing-impaired children

Colour-number accuracy

Figure 6.6A shows the percentages of trials on which participants correctly reported the Colour-number combination that was spoken by the target talker. A 3 x 2 repeated-

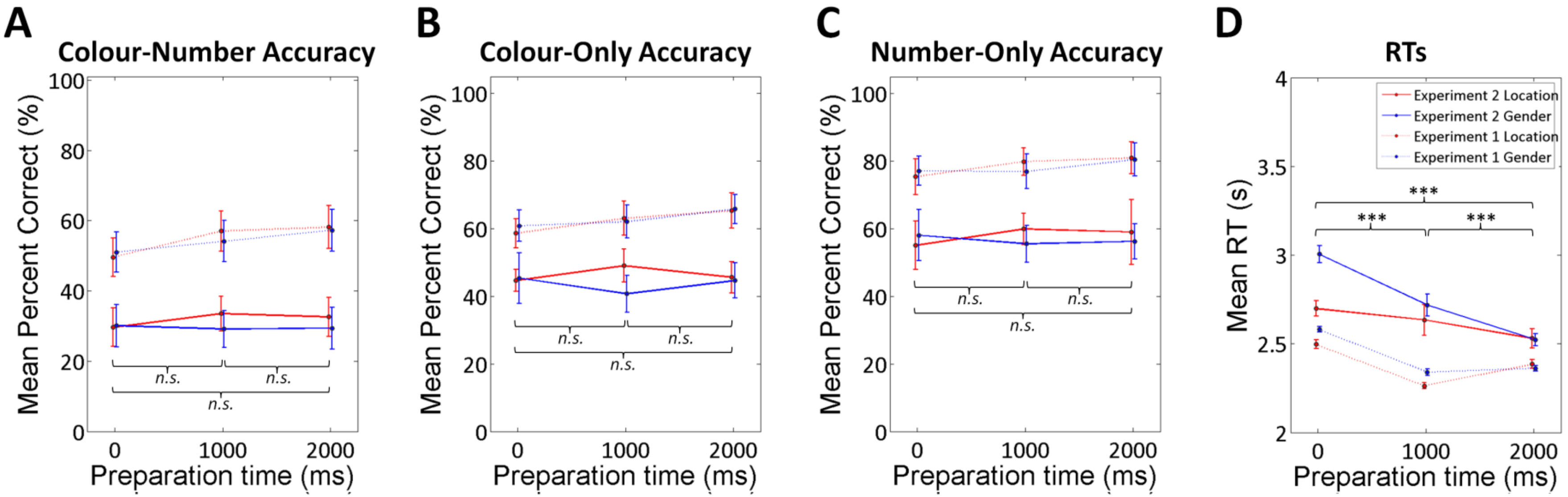


Figure 6.6. Accuracy and reaction time (RT) results for Experiment 2 (solid lines) with the results from Experiment 1 shown as a comparison (dashed lines). (**A**) Mean percentage of trials in which participants correctly identified the colour-number combination spoken by the target talker. (**B**) Mean percentage of trials in which participants correctly identified the colour (irrespective of the number). (**C**) Mean percentage of trials in which participants correctly identified the number (irrespective of the colour). (**D**) Mean RTs on trials in which participants correctly identified the colour-number combination. Error bars show within-subjects 95% confidence intervals. Brackets indicate the significance level of contrasts comparing 0-ms and 1000-ms and 0-ms and 2000-ms and post-hoc tests comparing 1000-ms and 2000-ms (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$) for the results from Experiment 2.

measures ANOVA showed no significant main effect of cue-target interval [F(1, 8) = 0.73, p = 0.42], no significant main effect of cue type [F(2, 16) = 0.38, p = 0.69], and no significant interaction [F(2, 16) = 0.62, p = 0.55].

Colour-only accuracy

The pattern for Colour-only accuracy (Figure 6.6B) was similar to the pattern for Colour-number accuracy. There were no significant main effects of cue-target interval [$F(1, 8) = 1.31$, $p = 0.29$] or cue type [$F(2, 16) = 0.01$, $p = 0.99$] and no significant interaction [$F(2, 16) = 2.77$, $p = 0.09$].

Number-only accuracy

Figure 6.6C shows the percentages of trials on which participants correctly reported the number that was spoken by the target talker. There were no significant main effects of cue-target interval [$F(1, 8) = 0.22$, $p = 0.65$] or cue type [$F(2, 16) = 0.13$, $p = 0.88$] and no significant interaction [$F(2, 16) = 0.86$, $p = 0.44$].

RTs

Even though *accuracy* did not change as a function of cue-target interval, RTs became shorter as the duration of the cue-target interval increased (Figure 6.6D). There was a significant main effect of cue-target interval, $F(1, 8) = 118.75$, $p < 0.001$, $\eta_p{}^2 = 0.94$. Contrasts showed that the 1000-ms [$F(1, 8) = 67.82$, $p < 0.001$, $\eta_p{}^2 = 0.82$] and 2000-ms [$F(1, 8) = 146.21$, $p < 0.001$, $\eta_p{}^2 = 0.95$] intervals both produced significantly shorter RTs than the 0-ms interval. Bonferroni-corrected post-hoc tests also showed significantly shorter RTs at the 2000-ms than the 1000-ms interval ($p < 0.001$).

RTs were significantly shorter on attend-location trials (M = 2.62 s, SD = 0.08) than attend-gender trials (M = 2.75 s, SD = 0.06), $F(2, 16) = 32.32$, $p < 0.001$, $\eta_p{}^2 = 0.80$. There was also a significant two-way interaction between cue-target interval and cue type, $F(1.2, 9.5) = 18.63$, $p = 0.001$, $\eta_p{}^2 = 0.70$. Bonferroni-corrected post-hoc tests showed significant differences between all cue-target intervals for attend-gender trials ($p \leq 0.001$), but only between 0-ms and 2000-ms for attend-location trials ($p = 0.002$).

Errors

There were similar percentages across all error types. The largest percentage of errors were "child" errors (M = 29.8%, SD = 8.1), followed by "opposite-gender" errors (M = 27.6%, SD = 11.6), "absent" errors (M = 25.5%, SD = 15.0), and "mix" errors (M = 17.2%, SD = 4.6). The percentages of "opposite-gender" [$t(8) = 5.39$, $p = 0.001$] and "child" [$t(8) = 8.73$, $p < 0.001$] errors were significantly greater than their expected values, whereas the percentages of "mix" [$t(8) = 14.95$, $p < 0.001$] and "absent" [$t(8) = 4.28$, $p = 0.003$] errors were significantly smaller than their expected values.

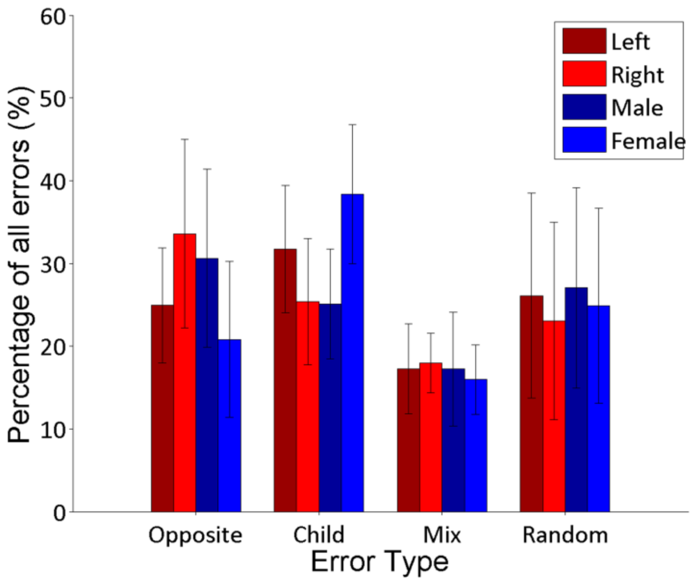


Figure 6.7. Experiment 2: Comparison of the percentages of incorrect responses that were categorised into each of the four error-types ("opposite-gender", "child", "mix", or "random" errors) for each of the four different cue types (left/right/male/female).

A 4 x 3 x 4 repeated-measures ANOVA investigated whether the percentages of different error types (4 levels: "opposite-gender", "child", "mix", and "absent" errors) differed significantly between different cue-target intervals (3 levels: 0, 1000, and 2000 ms) or between cue types (4 levels: left, right, male, and female). There was no significant main effect of error type [$F(1.7, 13.7) = 1.93$, $p = 0.19$] and no significant difference in the percentages of errors made for different cue-target intervals [error type * cue-target interval interaction: $F(6, 48) = 1.33$, $p = 0.74$]. However, there was a significant difference in the percentages of errors across the four different cue types [Figure 6.7; error type * cue type interaction: $F(9, 72) = 4.69$, $p < 0.001$, $\eta_p^2 = 0.37$]. When the cue-type variable was collapsed into attend-location (average left/right) and attend-gender (average male/female) trials, there were no significant differences in the percentages of errors made between attend-location and attend-gender trials overall [error type * cue type interaction: $F(3, 24) = 2.60$, $p = 0.08$].

6.4.2.2. Comparisons between Experiments 1 and 2

Accuracy and RTs

To compare the patterns of accuracy and RTs for different cue-target intervals and different cue-types between normally-hearing and hearing-impaired children, a 2 x 3 x 2 mixed ANOVA was conducted with the factors of hearing loss (2 levels: normally-hearing/hearing-impaired), cue-target interval (3 levels: 0/1000/2000 ms), and cue type (2 levels: location/gender). The analysis was conducted separately for Colour-number accuracy, Colour-only accuracy, Number-only accuracy, and RTs.

Colour-number accuracy was significantly higher for normally-hearing children (M = 54.5%, SD = 19.3) than hearing-impaired children (M = 30.7%, SD = 19.3), $F(1, 27) = 9.40$, $p =$

0.005, $\eta_p^2 = 0.26$. There was a significant main effect of cue-target interval [$F(2, 54) = 7.47$, $p = 0.001$, $\eta_p^2 = 0.22$] and a significant interaction between cue-target interval and hearing group [$F(2, 54) = 3.63$, $p = 0.033$, $\eta_p^2 = 0.12$]. This interaction reflects the result that increasing the duration of preparation time led to significantly higher colour-number accuracy for normally-hearing children, but not for hearing-impaired children. However, there was no significant main effect of cue type [$F(1, 27) = 0.21$, $p = 0.65$] and no significant interaction between cue type and hearing group [$F(1, 27) = 0.06$, $p = 0.81$]. The fact that this interaction was not significant reflects the finding that neither group showed significant variation in Colour-number accuracy between attend-location and attend-gender trials (Sections 6.3.2.1 and 6.4.2.1).

Colour-only accuracy was also significantly higher for normally-hearing children (M = 62.6%, SD = 14.7) than hearing-impaired children (M = 45.0%, SD = 14.7), $F(1, 27) = 8.93$, $p = 0.006$, $\eta_p^2 = 0.25$. There was a significant main effect of cue-target interval [$F(2, 54) = 3.58$, $p = 0.035$, $\eta_p^2 = 0.12$] and a significant interaction between cue-target interval and hearing group [$F(2, 54) = 3.30$, $p = 0.044$, $\eta_p^2 = 0.11$]. However, there was no significant main effect of cue type [$F(1, 27) = 0.17$, $p = 0.68$] and no significant interaction between cue type and hearing group [$F(1, 27) = 0.39$, $p = 0.54$].

Number-only accuracy was significantly higher for normally-hearing children (M = 78.5%, SD = 17.4) than hearing-impaired children (M = 57.3%, SD = 17.4), $F(1, 27) = 9.11$, $p = 0.005$, $\eta_p^2 = 0.25$. However, there was no significant main effect of cue-target interval [$F(2, 54) = 2.11$, $p = 0.13$], no significant interaction between cue-target interval and hearing group [$F(2, 54) = 0.85$, $p = 0.43$], no significant main effect of cue type [$F(1, 27) = 0.12$, $p = 0.73$], and no significant interaction between cue type and hearing group [$F(1, 27) = 0.02$, $p = 0.88$].

RTs were significantly shorter for normally-hearing children (M = 2.41 s, SD = 0.14) than hearing-impaired children (M = 2.69 s, SD = 0.14), $F(1, 27) = 25.0$, $p < 0.001$, $\eta_p^2 = 0.48$. There was a significant main effect of cue-target interval [$F(2, 54) = 418.42$, $p < 0.001$, $\eta_p^2 = 0.94$] and a significant interaction between cue-target interval and hearing group [$F(2, 54) = 78.27$, $p < 0.001$, $\eta_p^2 = 0.74$]. This interaction reflected differences in the pattern of RTs for normally-hearing and hearing-impaired children between 1000 and 2000 ms—normally-hearing children showed longer RTs at 2000 than 1000 ms, whereas hearing-impaired children showed shorter RTs. There was a significant main effect of cue type [$F(1, 27) = 63.90$, $p < 0.001$, $\eta_p^2 = 0.70$] and a significant interaction between cue type and hearing group [$F(1, 27) = 14.49$, $p = 0.001$, $\eta_p^2 = 0.35$]. This interaction reflected a larger difference in average RTs between attend-location and attend-gender trials in hearing-impaired children than normally-hearing children.

Errors

To compare how the percentages of different error types ("opposite-gender", "child", "mix", and "absent" errors) differed between normally-hearing and hearing-impaired children, a 2 x 4 mixed ANOVA was performed with the factors error type and hearing group. Since neither group showed significant effects of preparation time or cue type on the proportion of different error types, these factors were collapsed in the current analysis. The interaction between error type and hearing group was significant [$F(3, 81) = 3.56$, $p = 0.018$, $\eta_p^2 = 0.12$]. This result reflects similar percentages across all error types in hearing-impaired children but higher percentages of "opposite-gender" and "child" errors than "mix" and "absent" errors in normally-hearing children.

6.4.3. Discussion

Experiment 2 showed no effect of the duration of the cue-target interval on the accuracy of speech intelligibility in a group of nine mildly and moderately hearing-impaired children. However, there was a significant effect of preparation time on RTs: RTs were shorter when the cue-target interval was 2000 ms than when it was 0 ms for both attend-location and attend-gender trials. There was a significant progressive shortening of RTs in attend-gender trials from 0 to 1000 ms and from 1000 to 2000 ms, but this pattern was not present in attend-location trials. Nevertheless, RTs for both conditions were similar at the 2000-ms interval. Therefore, one possible reason for different patterns of results in attend-location and attend-gender trials is that preparation time provides a greater benefit when baseline RTs are longer. Overall, the results suggest that hearing-impaired children achieve shorter latencies for correctly reporting key words spoken by a target talker when they receive a cue for location or gender 2000 ms before compared to 0 ms before a target talker begins to speak.

The finding of shorter latencies at longer cue-target intervals suggests that hearing-impaired children engage in at least some preparatory processing. Since Experiment 1 found the opposite pattern of results—an improvement in accuracy but no improvement in latency with increasing durations of preparation time—the speed-accuracy trade-off might explain different results between normally-hearing and hearing-impaired children. Therefore, to confirm the finding that hearing-impaired children do not achieve improved accuracy of speech intelligibility with longer durations of preparation time, a replication would be desirable.

6.5. Experiment 3

Experiment 3 measured brain activity using electro-encephalography (EEG) in normally-hearing children during a similar three-talker listening task as was presented in Experiments 1 and 2. Experiment 1 showed that normally-hearing children achieved higher accuracy of speech intelligibility, whilst making responses with shorter latencies, when they had time to prepare for the location or gender of an upcoming talker compared to when they had no time to prepare. Therefore, it was expected that Experiment 3 would reveal preparatory EEG activity when normally-hearing children were cued to the location or gender of an upcoming talker. In addition, normally-hearing children were expected to display a similar pattern of ERPs as displayed by the normally-hearing adults reported in Chapter 5. This hypothesis followed from the results reported in Chapter 3, which showed similar patterns of ERPs for normally-hearing children and adults in a two-talker listening task.

6.5.1. Methods

6.5.1.1. Participants

Participants were 24 children (9 male), aged 8–15 years (M = 12.3, SD = 1.9). All participants were declared by their parents to be right-handed native English speakers with no history of hearing problems. They had 5-frequency average pure-tone hearing levels of 20 dB HL or better, tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). None of the children had previously taken part in any of the experiments reported in this thesis. The study was approved by the Research Ethics Committee of the Department of Psychology of the University of York.

6.5.1.2. Apparatus, stimuli, and procedure

The apparatus, stimuli, and procedure were identical to the experiment reported in Chapter 5. The average presentation level of concurrent triplets of test sentences was set to 63 dB(A) (range 61.6—66.2 dB) measured with a B&K (Brüel & Kjær, Nærum, Denmark) Sound Level Meter (Type 2260 Investigator) and 0.5-inch Free-field Microphone (Type 4189) placed in the centre of the arc at the height of the loudspeakers with the participant absent.

Figure 6.8 shows the trial structure of the Test and Control Conditions. All participants completed 96 trials in the Control Condition, 12 trials in the first set of familiarisation trials, 4 in the second set, and between 96 and 144 trials in the Test Condition (depending on their level of fatigue).

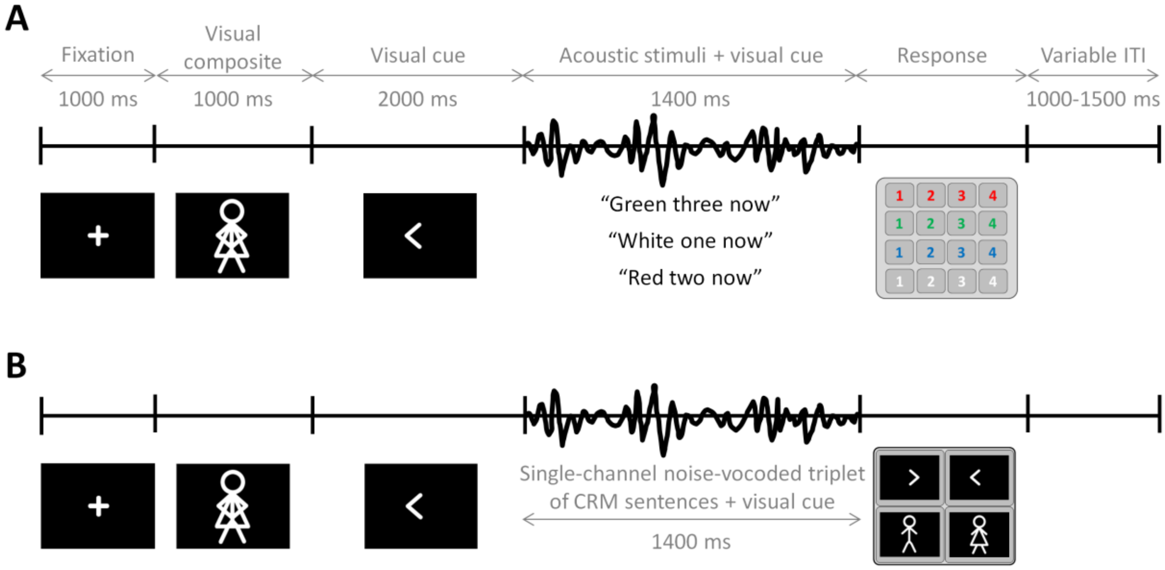


Figure 6.8. Trial structure of Experiments 3 and 4. (**A**) Trial structure in the Test Condition, with an example trial below. (**B**) Trial structure in the Control Condition.

6.5.1.3. EEG recording and processing

EEG recording and processing were identical to the experiments reported in Chapters 3 and 5 (described in Section 3.1.1.5).

6.5.1.4. Behavioural analyses

Trials were separated into Location (average left/right cues) and Gender (average male/female cues) groups, separately for the Test and Control Conditions. Responses were scored as correct if both the colour and number key words were reported correctly in the Test Condition, and if the visual cue was reported correctly in the Control Condition.

6.5.1.5. Analyses of ERPs

In seeking significant differences, a Spatio-temporal Cluster-based Permutation Analysis was performed. The analyses were identical to the analyses performed in Chapters 3 and 5. Type-I analyses compared the Test and Control conditions, separately for Location and Gender trials. Type-II analyses compared Location and Gender trials in the Test Condition. Due to differences in physical attributes of the visual cues between Location and Gender Conditions, for each Type-II cluster, the average amplitude of Location and Gender trials—averaged over the space-by-time points in the cluster—was compared between the Test and Control Conditions in a 2 x 2 repeated-measures ANOVA. A two-way interaction meant that the cluster could not be fully explained by the influence of physical differences in the visual cues between conditions. In order to determine whether such differences were sustained over the entire duration of a cluster or were restricted to particular moments, we plotted the difference of the difference in Location and Gender trials between the Test and Control Conditions, averaging only over the space-by-time points that fell in a 50-ms time window that was advanced in 10-ms steps over the duration of the cluster.

6.5.2. Results

6.5.2.1. Behavioural results

Conjoint accuracy in identifying the colour and number key words in the Test Condition differed significantly between Location (M = 68.5%, SD = 12.9) and Gender (M = 64.1%, SD = 16.6) trials, *t*(23) = 2.35, *p* = 0.028. There were no significant differences in the accuracy with which the visual cue was identified in the Control Condition between Location (M = 97.7%, SD = 4.2) and Gender (M = 98.5%, SD = 2.6) trials, *t*(23) = 0.88, *p* = 0.39.

6.5.2.2. Event-related potentials

Type-I analyses: Differences between Test and Control Conditions

Location trials

Figure 6.9 illustrates the results of the Type-I analyses on trials in which a Location cue (left/right) was presented. During the 2000-ms Preparatory Phase, three significant clusters of activity were identified (Figure 6.9A). Cluster 1 involved 26 posterior electrodes and spanned the time interval from 74 to 370 ms, relative to the start of the phase. Cluster 1 showed significantly more negative amplitude in the Test Condition than the Control Condition (cluster statistic = 11653, *p* = 0.040; Figure 6.9B). The polarity, location, onset time, and duration of Cluster 1 are tabulated in the first line of the first column of Table 6.1. Cluster 2 (Figure 6.9C) spanned the interval from 830 to 2000 ms. It involved 32 posterior and central electrodes and showed significantly more negative amplitude in the Test Condition than the Control Condition [cluster statistic = 49926, *p* < 0.001]. Cluster 3 (Figure 6.9D; 1035 to 2000 ms) was complementary to Cluster 2 because it overlapped in time with Cluster 2 but spanned a different group of electrodes with opposite polarity. Cluster 3 involved 28 mainly anterior electrodes and showed significantly more positive amplitude in the Test Condition than the Control Condition [cluster statistic = 28580, *p* = 0.003]. Clusters 1–3 demonstrate that differences in brain activity arise between a condition in which a visual cue has no implications for auditory attention and a condition in which the same visual cue directs listeners to prepare for the location of an upcoming talker.

During the Selective Phase, three significant clusters of activity were identified (Figure 6.9A). Cluster 4 (Figure 6.9E) involved 31 anterior electrodes and spanned the time interval from 0 to 556 ms, relative to the start of the phase. Cluster 4 showed significantly

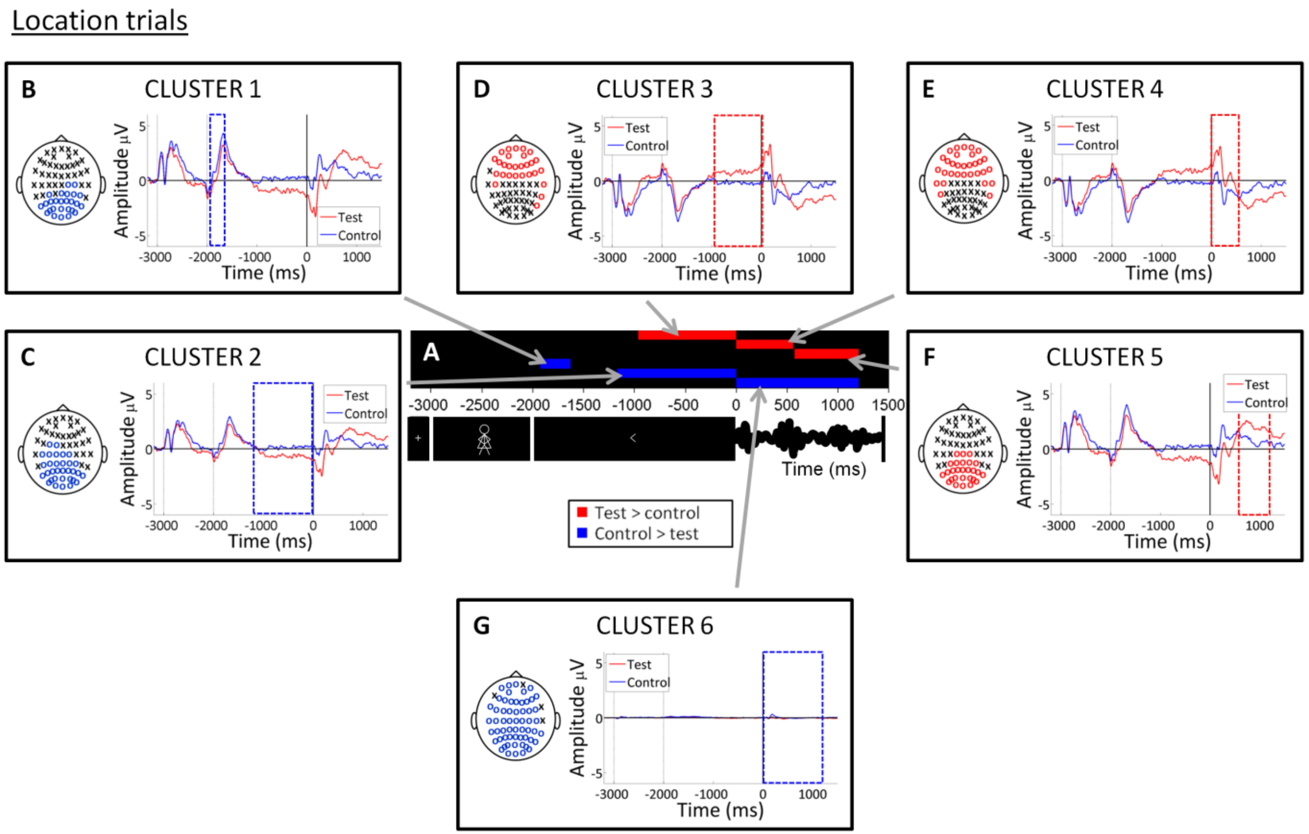


Figure 6.9. (Continued on next page). Experiment 3: Results from Type-I Spatio-Temporal Cluster-based Permutation Analyses for the Location Condition (**A** to **G**) and the Gender Condition (**H** to **N**). (**A** and **H**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Further information about each cluster is displayed in (**B**)-(**G**) and (**J**)-(**N**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

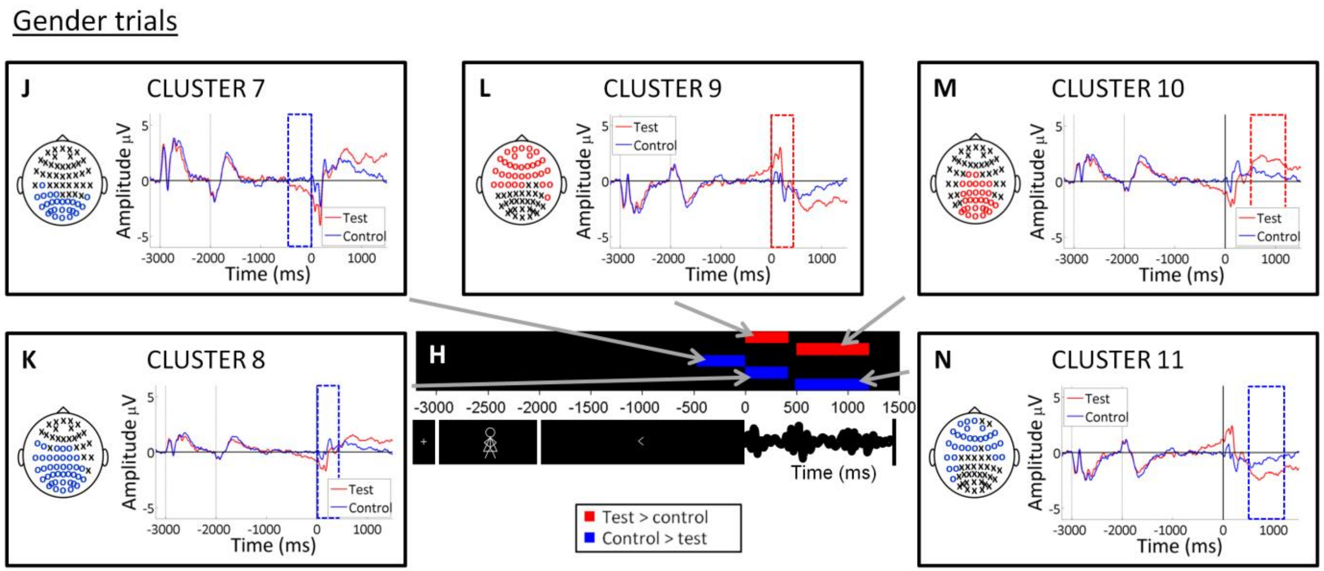


Figure 6.9. (Continued from the previous page)

Table 6.1. (Continued on next page). Experiments 3 and 4: Summary of comparisons in Gender and Location trials (Type-I analysis) between the Test and Control Conditions. The results for Experiment 3 are reported for correct-only trials, whereas the results for Experiment 4 are reported for correct-and-incorrect trials. (The results of the Cluster-based Permutation Analyses for correct-only trials in Experiment 4 are displayed in Appendix 3).

Phase		**Experiment 3 Location**	**Experiment 4 Location**	**Experiment 3 Gender**	**Experiment 4 Gender**
Preparatory	**Cluster Number**	**1**	**-**	**-**	**15**
	Cluster *p*-value	0.040	-	-	0.029
	Polarity	Control > Test	-	-	Control > Test
	Electrode Locations	Posterior	-	-	Central + Posterior
	Onset of cluster (ms)	74	-	-	0
	Duration of cluster (ms)	296	-	-	452
Preparatory	**Cluster Number**	**2**	**-**	**7**	**-**
	Cluster *p*-value	< 0.001	-	0.024	-
	Polarity	Control > Test	-	Control > Test	-
	Electrode Locations	Central + Posterior	-	Posterior	-
	Onset of cluster (ms)	830	-	1527	-
	Duration of cluster (ms)	1170	-	473	-
Preparatory	**Cluster Number**	**3**	**-**	**-**	**-**
	Cluster *p*-value	0.003	-	-	-
	Polarity	Test > Control	-	-	-
	Electrode Locations	Anterior	-	-	-
	Onset of cluster (ms)	1035	-	-	-
	Duration of cluster (ms)	965	-	-	-
Selective *(continued on next page)*	**Cluster Number**	**4**	**-**	**9**	**-**
	Cluster *p*-value	< 0.001	-	0.005	-
	Polarity	Test > Control	-	Test > Control	-
	Electrode Locations	Anterior	-	Central + Anterior	-

Table 6.1. (Continued from previous page)

Phase		Experiment 3 Location	Experiment 4 Location	Experiment 3 Gender	Experiment 4 Gender
Selective *(continued from previous page)*	**Onset of cluster (ms)**	0	-	0	-
	Duration of cluster (ms)	556	-	423	-
Selective	**Cluster Number**	**5**	**-**	**10**	**-**
	Cluster *p*-value	< 0.001	-	< 0.001	-
	Polarity	Test > Control	-	Test > Control	-
	Electrode Locations	Posterior	-	Central + Posterior	-
	Onset of cluster (ms)	577	-	495	-
	Duration of cluster (ms)	623	-	705	-
Selective	**Cluster Number**	**6**	**14**	**8**	**16**
	Cluster *p*-value	< 0.001	0.020	0.001	0.001
	Polarity	Control > Test	Control > Test	Control > Test	Control > Test
	Electrode Locations	Anterior + Central + Posterior	Central + Posterior	Central + Posterior	Central + Posterior
	Onset of cluster (ms)	0	12	0	12
	Duration of cluster (ms)	1200	469	416	531
Selective	**Cluster Number**	**-**	**-**	**11**	**17**
	Cluster *p*-value	-	-	< 0.001	0.033
	Polarity	-	-	Control > Test	Control > Test
	Electrode Locations	-	-	Anterior	Anterior
	Onset of cluster (ms)	-	-	485	910
	Duration of cluster (ms)	-	-	715	290

more positive amplitude in the Test Condition than the Control Condition [cluster statistic = 32568, $p < 0.001$]. Cluster 5 (Figure 6.9F) spanned the interval from 577 to 1200 ms. It involved 27 posterior electrodes and showed significantly more positive amplitude in the Test Condition than the Control Condition [cluster statistic = 39909, $p < 0.001$]. Cluster 6

(Figure 6.9G) spanned the entire Selective Phase (0 to 1200 ms, relative to the start of the phase). Cluster 6 involved 58 electrodes across almost the entire electrode array and showed significantly more negative amplitude in the Test Condition than the Control Condition [cluster statistic = 65681, $p < 0.001$].

Gender trials

In the second of the Type-I analyses, ERPs between the Test and Control Conditions were compared on trials in which a Gender cue (male/female) was presented. Panels H–N of Figure 6.9 show these results. One significant cluster was identified during the Preparatory Phase. Cluster 7 occurred towards the end of the Preparatory Phase. It involved 25 posterior electrodes and spanned the time interval from 1527 to 2000 ms, relative to the start of the phase. Cluster 7 showed significantly more negative amplitude in the Test Condition than the Control Condition [cluster statistic = 13119, $p = 0.024$] (Figure 6.9J). Cluster 7 demonstrates a difference in brain activity between a condition in which a visual cue has no implications for auditory attention and in which the same visual cue directs listeners to prepare for the gender of an upcoming talker.

During the Selective Phase, four significant clusters were identified (Figure 6.9H). Cluster 8 (Figure 6.9K) involved 39 central and posterior electrodes and spanned the time interval from 0 to 416 ms relative to the start of the phase. It showed significantly more negative amplitude in the Test Condition than the Control Condition [cluster statistic = 29011, $p = 0.001$]. Cluster 9 (Figure 6.9L; 0 to 423 ms) was complementary to Cluster 8. Cluster 9 involved 31 central and anterior electrodes and showed significantly more positive amplitude in the Test Condition than the Control Condition [cluster statistic = 21405, $p = 0.005$]. Cluster 10 (Figure 6.9M) involved 29 central and posterior electrodes and spanned the time interval from 495 to 1200 ms. It showed significantly more positive amplitude in the Test Condition than the Control Condition [cluster statistic = 41491, $p < 0.001$]. Cluster 11 (Figure 6.9M; 485 to 1200 ms) was complementary to Cluster 10. Cluster 11 involved 31 mainly anterior electrodes and showed significantly more negative amplitude in the Test Condition than the Control Condition [cluster statistic = 31541, $p < 0.001$].

Type-II analyses: Differences between Location and Gender trials

Differences during the Preparatory Phase

During the Preparatory Phase, one cluster of activity was identified that differed significantly in the Test Condition between Location and Gender trials (Figure 6.10A). Cluster 12 (Figure 6.10B; 943 ms to 1604 ms) involved 19 anterior electrodes and showed significantly more positive amplitude during Location than Gender trials [cluster statistic =

Table 6.2. Experiment 3: Summary of results for the Test Condition comparison between Location and Gender trials (Type-II analysis). A tick in the row headed ‘Significant in Control Condition?’ indicates that the difference in the amplitude of ERPs between Location and Gender trials was significant in the Control Condition across the spatio-temporal points of the cluster (p-values displayed underneath). A tick in the row headed ‘Significant Test/Control Interaction?’ indicates that an ANOVA with the factors cue type (Location/Gender) and condition (Test/Control) revealed a significant two-way interaction (p-values displayed underneath).

Phase	Properties	Cluster Properties
Preparatory	**Cluster Number**	**12**
	Cluster p-value	0.018
	Polarity	Location > Gender
	Electrode Locations	Anterior
	Onset of cluster (ms)	943
	Duration of cluster (ms)	661
	Significant in Control Condition?	✗ p = 0.92
	Significant Test/Control Interaction?	✓ p = 0.008
Selective	**Cluster Number**	**13**
	Cluster p-value	0.004
	Polarity	Gender > Location
	Electrode Locations	Central + Central-Posterior
	Onset of cluster (ms)	365
	Duration of cluster (ms)	644
	Significant in Control Condition?	✗ p = 0.88
	Significant Test/Control Interaction?	✓ p = 0.002

12408, p = 0.018]. These values are listed in the first line of Table 6.2. For this cluster, the interaction between cue type (Location/Gender) and condition (Test/Control) was significant [$F(1,23) = 8.57$, p = 0.008, $\eta_p^2 = 0.27$; Figure 6.11A] and the difference between Location and Gender trials was not significant in the Control Condition, p = 0.92. When the difference of the differences in Location and Gender trials between the Test and Control conditions was examined in 50-ms sliding windows, the uncorrected p-value was less than 0.05 in 11 of the 62 50-ms time windows in the cluster (Figure 6.11C). The finding that ERPs in this cluster differed significantly between the Test and Control Conditions means that it is possible to rule out the explanation that the cluster arose from differences in the visual cues. Instead, the cluster must reflect differential attentional preparation for an upcoming talker based on cues for location and gender.

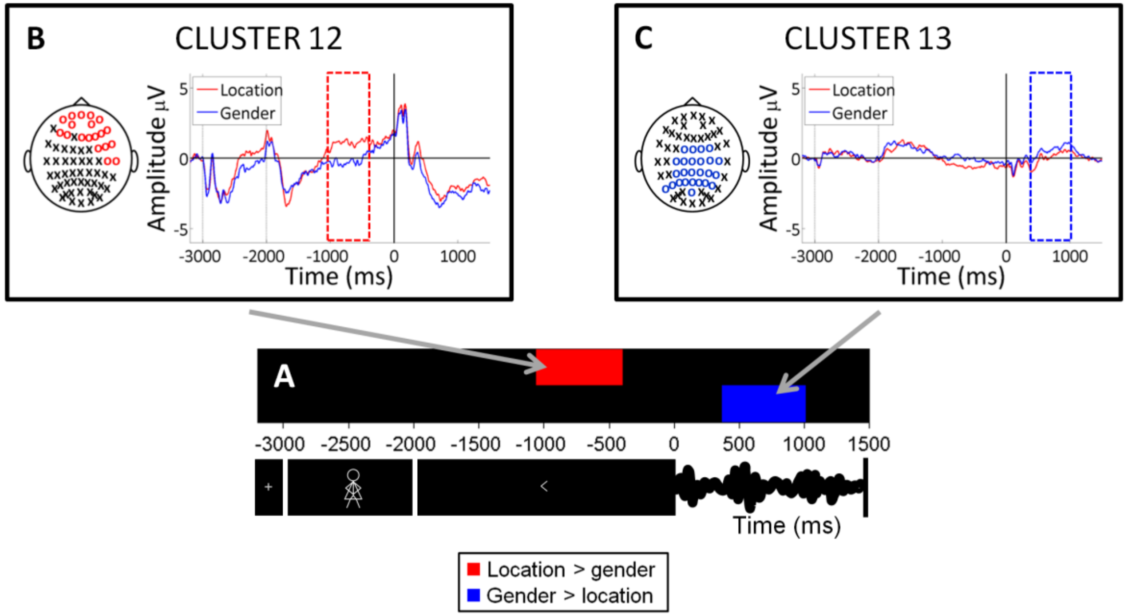


Figure 6.10. Experiment 3: Results of the Type-II Spatio-Temporal Cluster-based Permutation Analysis. This analysis contrasted Location and Gender trials in the Test Condition. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive on location trials than gender trials. For clusters plotted as blue rectangles, the average amplitude was more negative on location trials than gender trials. Further information about each cluster is displayed in (**B**)-(**C**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

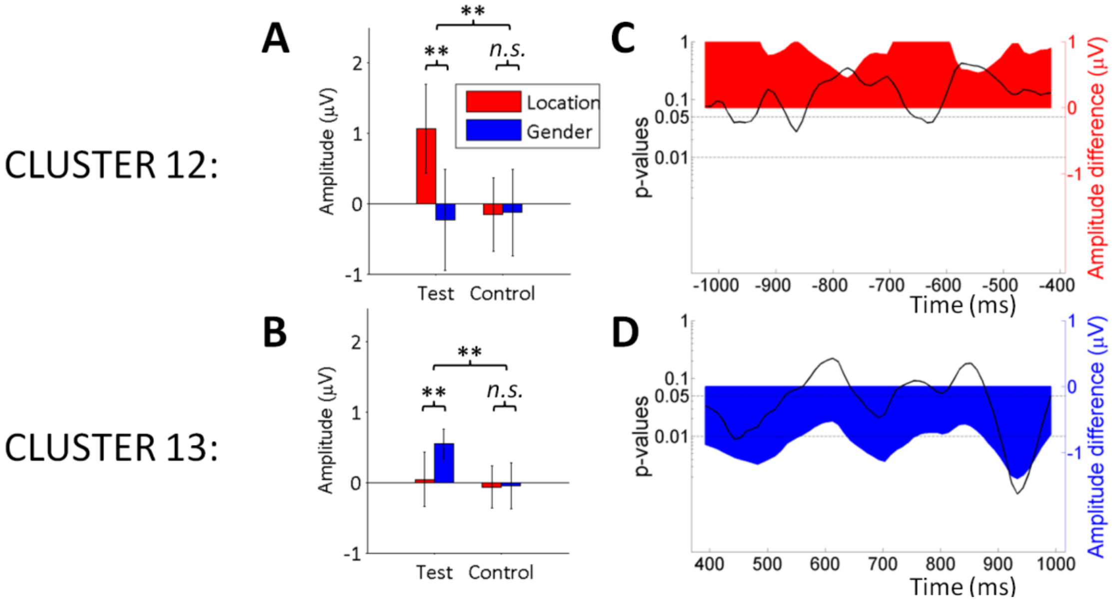


Figure 6.11. Experiment 3: Comparison of differences in the amplitude of ERPs between Location and Gender trials in the Test and Control Conditions for each significant Type-II cluster. Graphs (**A**)-(**B**) plot the mean amplitude for Location and Gender trials in the Test and Control Conditions, averaged across participants and space-time points. Error bars show 95% within-subjects confidence intervals. Narrow brackets display the significance level of the comparison between Location and Gender trials in the Test and Control Conditions. Wider brackets display the significance level of the two-way interaction (* $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). Graphs (**C**)-(**D**) display the difference of the differences in Gender and Location trials between the Test and Control conditions in 50-ms time windows repeated every 10 ms within the cluster (right axis) and the uncorrected *p*-values resulting from a paired-samples *t*-test comparing the differences (left axis). The mid-point of each time window relative to the onset of acoustic stimuli is displayed on the x-axis.

Differences during the Selective Phase

During the Selective Phase, one cluster of activity was identified that differed significantly between Location and Gender trials (Figure 6.10A). Cluster 13 (Figure 6.10C) lasted from 365 to 1009 ms after the start of the Selective Phase. It involved 25 central and central-posterior electrodes and displayed significantly more negative amplitude in Location trials than Gender trials [cluster statistic = 12701, $p = 0.004$]. The interaction between cue type (Location/Gender) and condition (Test/Control) was significant [$F(1,23) = 12.57$, $p = 0.002$, $\eta_p^2 = 0.35$; Figure 6.11B] and the difference between Location and Gender trials was not significant in the Control Condition, $p = 0.88$. The uncorrected *p*-value for the difference between the Test and Control Conditions was below 0.05 in 32 out of the 61 50-ms windows, which occurred at the beginning, middle, and end of the cluster (Figure 6.11D). Like Cluster

12, these findings demonstrate that Cluster 13 could not be explained by physical differences in the visual cues.

6.5.3. Discussion

Normally-hearing children showed significant ERPs during the Preparatory Phase depending on whether a visual cue directed attention to an upcoming talker or whether the same visual cues were presented but did not have implications for auditory attention. From these results, it can be inferred that the difference reflects attentional preparation for the location or gender of an upcoming target talker. Preparation for *location* evoked significant activity early after the visual cue was revealed. Preparation for *both* location and gender evoked later sustained preparation that occurred in the 500 ms immediately before the talkers started to speak.

In addition to similar activity, there were significant differences in EEG activity during the Preparatory Phase depending on whether participants were cued to the location or the gender of the target talker. This activity was not present in the Control Condition and, therefore, can be confidently attributed to cue-specific attentional processing based on knowledge of a target talker's location or gender. During the Preparatory Phase, cue-specific attentional preparation occurred between 500 and 1000 ms before the talkers began. This finding demonstrates that the mechanisms participants use to prepare for a target talker depends on knowledge of the talker's location or their gender. Cue-specific activity also occurred between 350 and 650 ms after the start of the Selective Phase, while the colour and number key words were spoken. This finding provides evidence that the brain activity involved in selecting talkers according to their location differs, in part, from the activity involved in selecting talkers according to their gender.

Overall, the results provide evidence for domain-general and cue-specific EEG activity. This pattern of results is similar to those reported for adults in Chapter 5 during an equivalent three-talker task.

6.6. Experiment 4

Experiment 4 measured EEG activity in hearing-impaired children during three-talker listening. The task was the same as that used in Experiment 3, except that the average presentation level of the acoustical stimuli was increased to 76 dB(A) SPL to ensure that the stimuli were audible for participants with bilateral moderate hearing loss. The presentation level was identical to that used in Experiment 2.

Experiment 2 found that hearing-impaired children were able to use longer durations of preparation time to improve the latency of correctly reporting words spoken by a target talker. This finding suggests that hearing-impaired listeners are able to utilise preparatory attention to some extent and, therefore, it was expected that the hearing-impaired listeners in Experiment 4 would show significant preparatory EEG activity. However, given that hearing-impaired children were slower and less accurate at reporting words spoken by a target talker than normally-hearing children (Experiments 1 and 2), hearing-impaired children were also expected to show different EEG activity to the normally-hearing children who participated in Experiment 3. Differences between normally-hearing and hearing-impaired children during the Selective Phase were expected to reflect atypical peripheral transduction in hearing-impaired children. Differences during the Preparatory Phase were expected to reflect differences in preparatory attention.

One method by which the ERPs were directly compared between normally-hearing and hearing-impaired children was by analysing the amplitudes that occurred in hearing-impaired children at the same space-by-time points at which significant differences were found in normally-hearing children in Experiment 3. Figure 6.12 displays the conclusions that were drawn from different patterns of evidence when comparing amplitudes in the Test and Control Conditions between normally-hearing and hearing-impaired children.

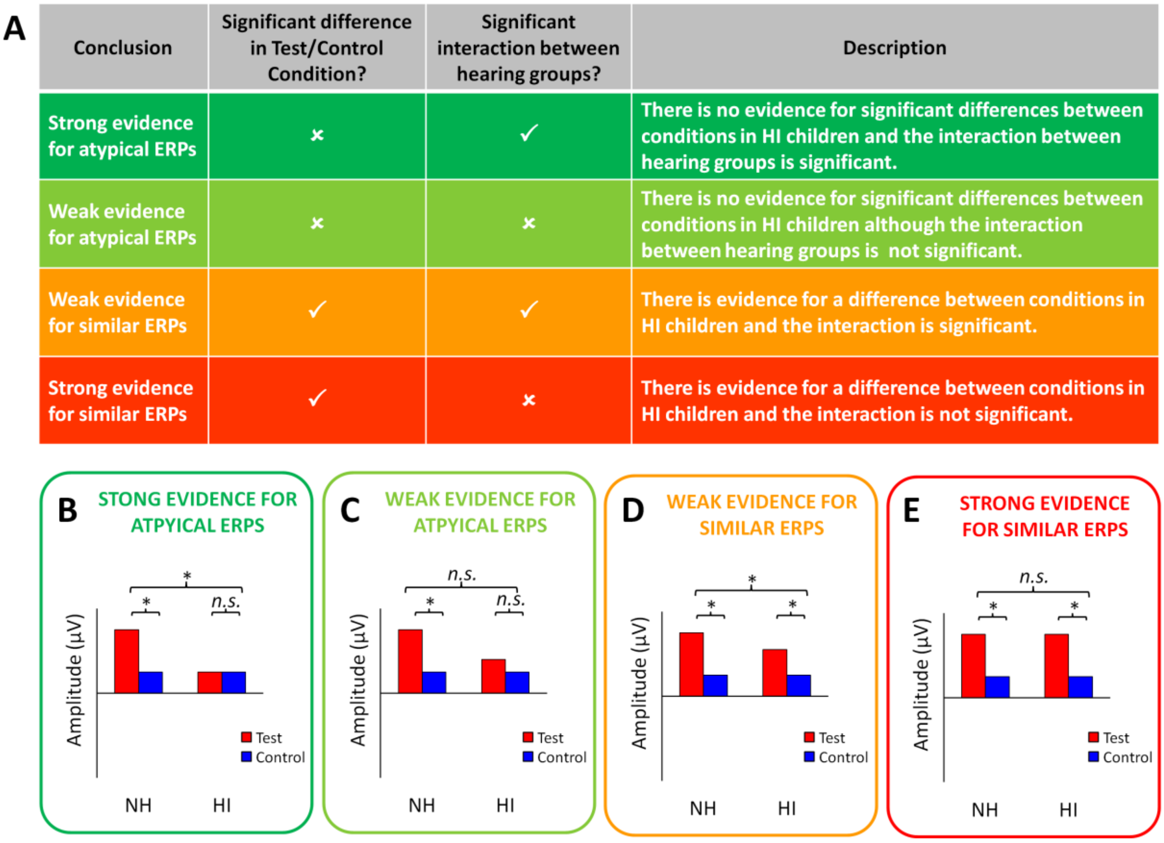
6.6.1. Methods

6.6.1.1. Participants

Participants were 13 children (3 male), aged 7–16 years (M = 11.9, SD = 3.0). Twelve children had moderate hearing loss and one child had mild hearing loss. They had bilateral 5-frequency average pure-tone hearing levels between 42 and 65 dB HL (M = 50.9 dB HL, SD = 8.0; Figure 6.13), tested in accordance with BS EN ISO 8253-1 (British Society of Audiology, 2004). The difference in the 5-frequency averages recorded from the left and right ears was less than 12 dB for each participant. Participants were declared by their parents to be native English speakers. Out of the thirteen children, one was left-handed and had an additional visual impairment in her left eye. Prior to taking part in this experiment, none of the children had taken part in any of the other experiments in this thesis[8]. Three additional children were tested, but had severe hearing loss and had difficulty identifying words spoken each of the talkers presented individually in quiet—therefore, they did not complete the experiment and their results are not included. The study was approved by the Research

---

[8] A subset of these children also participated in Experiment 2. However, they all participated in the current experiment without their hearing aids before they participated in Experiment 2.



A

Conclusion	Significant difference in Test/Control Condition?	Significant interaction between hearing groups?	Description
Strong evidence for atypical ERPs	×	✓	There is no evidence for significant differences between conditions in HI children and the interaction between hearing groups is significant.
Weak evidence for atypical ERPs	×	×	There is no evidence for significant differences between conditions in HI children although the interaction between hearing groups is not significant.
Weak evidence for similar ERPs	✓	✓	There is evidence for a difference between conditions in HI children and the interaction is significant.
Strong evidence for similar ERPs	✓	×	There is evidence for a difference between conditions in HI children and the interaction is not significant.

Figure 6.12. Possible conclusions from different patterns of ERP results when the average amplitudes of ERPs are compared between normally-hearing (NH) and hearing-impaired (HI) children for the clusters identified in NH children. Panel (**A**) tabulates the types of evidence that underlie different conclusions. A tick in the column headed 'Significant difference between Test/Control Conditions?' indicates that a paired-samples *t*-test revealed a significant difference in amplitude between the Test and Control Conditions in HI children within one cluster from NH children. A tick in the column headed 'Significant interaction between hearing groups?' indicates that a 2 x 2 between-subjects ANOVA with the factors hearing group (NH/HI) and condition (Test/Control) revealed a significant two-way interaction (*p*-values displayed underneath). The rows indicate different conclusions that would be drawn about the ERPs in HI children. The conclusion is related to the strength of the evidence (strong/weak) and the nature of the evidence (atypical ERPs in HI children or similar ERPs in NH and HI children). Panels (**B**)-(**D**) plot bar graphs showing hypothetical patterns of ERPs that could provide evidence to support different conclusions. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing the Test and Control Conditions within each hearing group. The difference between the Test and Control Conditions is expected to be significant in NH children, since the amplitudes are averaged over the space-by-time windows where significant clusters were observed in NH children. The larger brackets at the top of each graph indicate whether interaction between hearing group and the Test/Control Condition is significant.

Ethics Committee of the Department of Psychology, University of York, the NHS Research Ethics Committee of Newcastle and North Tyneside, and the Research and Development Departments of York Teaching Hospital NHS Foundation Trust, Leeds Teaching Hospitals NHS Trust, Hull and East Yorkshire Hospitals NHS Trust, and Bradford Teaching Hospitals NHS Foundation Trust.

Participants completed the experiment for the first time without using their hearing aids. A subset of nine hearing-impaired children also took part in the experiment for a second time using their own acoustic bilateral behind-the-ear hearing aids[9]. The results reported in the current chapter are from the first session in which listeners completed the task without their hearing aids. Results from the aided condition are reported in Appendix C.

6.6.1.2. Apparatus, stimuli, procedure, and EEG recording

The apparatus, stimuli, and procedure were the same as Experiment 3, except for the presentation level of acoustical stimuli. The average presentation level of concurrent triplets of test sentences was set to 76 dB(A) (range 72.4—77.9 dB) measured with a B&K (Brüel & Kjær, Nærum, Denmark) Sound Level Meter (Type 2260 Investigator) and 0.5-inch Free-field Microphone (Type 4189) placed in the centre of the arc at the height of the loudspeakers with the participant absent. Participants completed 96 trials in the Control Condition, 12 trials in the first set of familiarisation trials, 4 in the second set, and between 96 and 288 trials in the Test Condition. EEG recording and processing were the same as Experiment 3, except that incorrect trials were included in the analyses to maintain a high signal-to-noise ratio in participant averages (referred to as ‘correct-and-incorrect trials’). However, additional analyses reported in Appendix D also compared an analysis of correct-only trials (from which incorrect trials were excluded) with an analysis of correct-and-incorrect trials.

6.6.1.3. Behavioural analyses

Trials were separated into Location (average left/right cues) and Gender (average male/female cues) groups, separately for the Test and Control Conditions. Responses were scored as correct if both the colour and number key words were reported correctly in the Test Condition, and if the visual cue was reported correctly in the Control Condition.

6.6.1.4. Analyses of ERPs

Identical to Experiment 3, Spatio-temporal Cluster-based Permutation Analyses were used to perform Type-I and Type-II comparisons on the ERPs recorded from hearing-impaired children.

---

[9] The subset of children who participated in the aided condition of the current experiment was the same subset that participated in Experiment 2. They all participated in the aided condition after they had participated in Experiment 2.

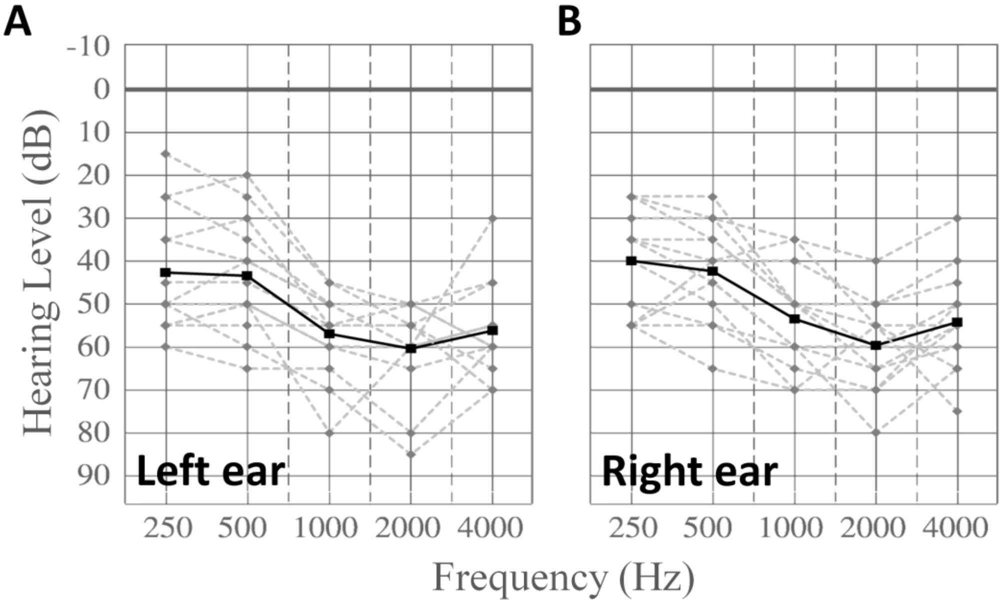


Figure 6.13. (above) Experiment 4: Pure-tone audiometric thresholds (dB HL) for each participant (grey dashed lines) and mean threshold across all participants (black solid line), plotted separately for the left (**A**) and right (**B**) ears.

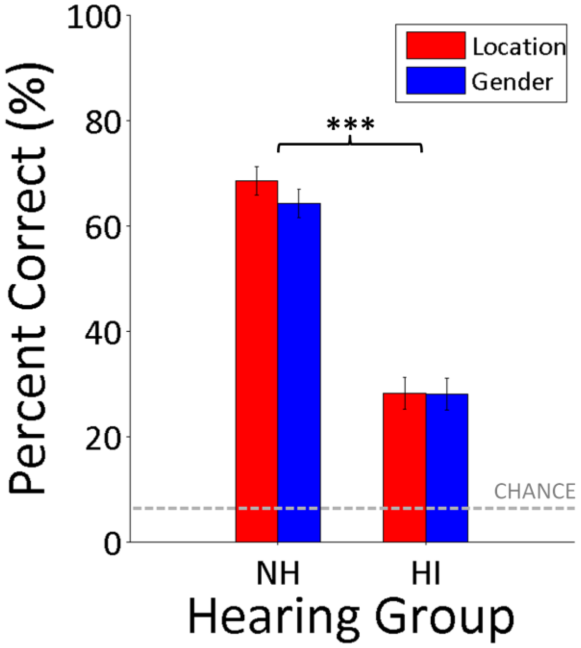


Figure 6.14. (left) Experiments 3 and 4. Mean percentage of trials in which participants correctly identified the colour-number combination spoken by the target talker, plotted as separate bars for the normally-hearing (NH) children in Experiment 3 and the hearing-impaired children (HI) in Experiment 4 who performed the task without their hearing aids.

6.6.1.5. Comparisons between Experiments 3 and 4

The behavioural and ERP results from the hearing-impaired children were also compared directly with the results from the normally-hearing children who participated in Experiment 3. The behavioural analysis compared accuracy between normally-hearing and hearing-impaired children.

The first ERP analysis compared the overall ERP waveforms measured from normally-hearing and hearing-impaired participants. First, trials were separated into those that occurred during the Test and Control Conditions. Next, amplitudes were averaged across a broad group of 28 posterior electrodes and a separate group of 34 anterior electrodes.

The second ERP analysis compared the average amplitude of activity in each cluster from normally-hearing children in Experiment 3—averaged over the space-by-time points in the cluster—between normally-hearing and hearing-impaired children. A summary of all of the analyses for Experiment 4 is shown in Table 6.3.

Table 6.3. Summary of the analyses reported in Experiment 4.

Comparison	Analyses	Section in which the results are reported
Within-subjects comparisons for hearing-impaired children in the aided session	Behavioural comparisons of Location and Gender trials	Section 6.6.2.1
	Type-I and Type-II ERP analyses	Section 6.6.2.2
Within-subjects analysis for hearing-impaired children between aided and unaided sessions	Behavioural comparisons of Location and Gender trials	Appendix C
	Overall ERP waveform analysis	Appendix C
	Type-I and Type-II ERP analyses	Appendix C
Within-subjects comparisons in hearing-impaired children unaided between correct-only trials and correct-and-incorrect trials	Behavioural comparisons of Location and Gender trials	Appendix D
	Overall ERP waveform analysis	Appendix D
	Type-I and Type-II ERP analyses	Section 6.6.2.2.3
Between-subjects comparisons of normally-hearing (correct-only trials) and hearing-impaired children unaided (correct-and-incorrect trials)	Behavioural comparisons of Location and Gender trials	Section 6.6.2.3
	Overall ERP waveform analysis	Section 6.6.2.3
	Type-I and Type-II ERP analyses	Section 6.6.2.3

6.6.2. Results

6.6.2.1. Behavioural results in hearing-impaired children

When participants completed the task unaided, conjoint accuracy in identifying the colour and number key words in the Test Condition did not differ significantly between Location (M = 28.2%, SD = 19.3) and Gender (M = 28.0%, SD = 17.1) trials, *t*(12) = 0.09, *p* = 0.93 (Figure 6.14). There were no significant differences in the accuracy with which the visual cue was identified in the Control Condition between Location (M = 92.8%, SD = 13.2) and Gender (M = 96.6%, SD = 4.6) trials, *t*(12) = 1.19, *p* = 0.26.

6.6.2.2. Event-related potentials in hearing-impaired children

Type-I analyses: Differences between Test and Control Conditions

Location trials

Figure 6.15 illustrates the results of the Type-I analyses on trials in which a Location cue (left/right) was presented. During the 2000-ms Preparatory Phase, no significant clusters of activity were identified. During the Selective Phase, one significant cluster of activity was identified in hearing-impaired children (Figure 6.15A). Cluster 14 (Figure 6.15B) involved 28 central and posterior electrodes and spanned the time interval from 12 to 481 ms, relative to the start of the phase. Cluster 14 showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 14048, *p* = 0.020]. The polarity, location, onset time, and duration of Cluster 1 are tabulated in the first line of the second column of Table 6.1.

Gender trials

In the second of the Type-I analyses, ERPs between the Test and Control Conditions were compared on trials in which a Gender cue (male/female) was presented. Panels C–F of Figure 6.15 show these results. One significant cluster was identified during the Preparatory Phase. Cluster 15 occurred towards the beginning of the Preparatory Phase. It involved 29 mainly central and posterior electrodes and spanned the time interval from 0 to 452 ms, relative to the start of the phase. Cluster 15 showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 8078, *p* = 0.029] (Figure 6.15D).

During the Selective Phase, two significant clusters were identified (Figure 6.15C). Cluster 16 (Figure 6.15E) involved 32 central and posterior electrodes and spanned the time interval from 12 to 543 ms relative to the start of the phase. It showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 20870, *p* = 0.001]. Cluster 17 (Figure 6.15F; 910 to 1200 ms) occurred towards the end of the Selective Phase. Cluster 17 involved 20 mainly anterior electrodes and showed significantly

Location trials

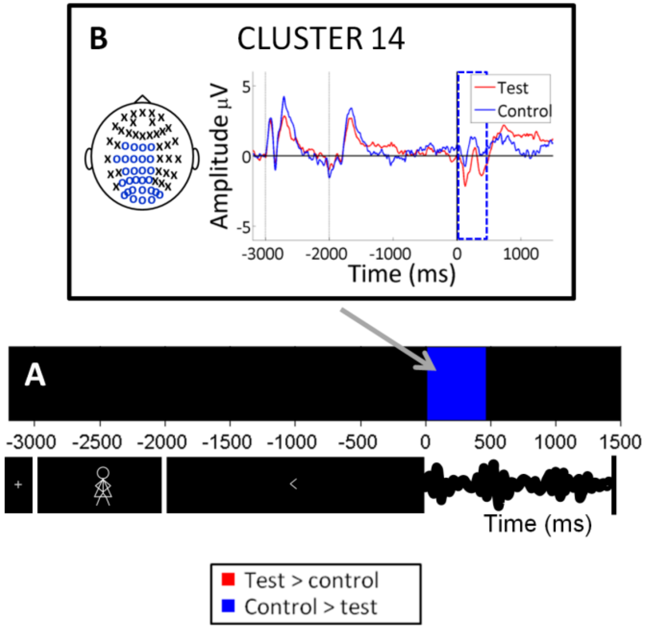


Figure 6.15. (Continued on the next page). Experiment 4: Results from Type-I Spatio-Temporal Cluster-based Permutation Analyses for Location (**A** to **B**) and Gender (**C** to **F**) trials. This analysis included correct-and-incorrect trials. (**A** and **C**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Further information about each cluster is displayed in (**B**) and (**D**)-(**F**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

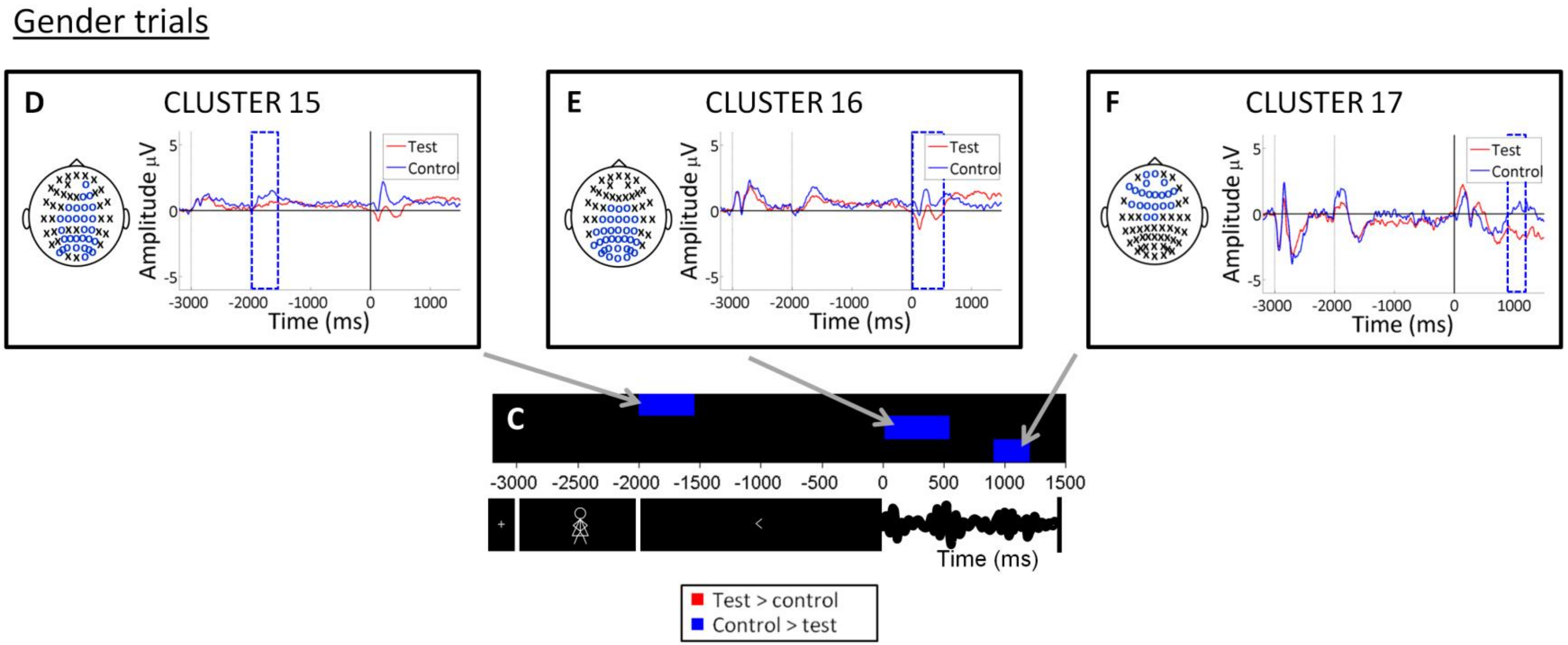


Figure 6.15. (Continued from the previous page)

more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 9655, p = 0.033].

Type-II analyses: Differences between Location and Gender trials

The Type-II analyses did not identify any significant clusters that differed in the Test Condition between Location and Gender trials, either during the Preparatory or Selective Phase.

6.6.2.3. Comparisons between Experiments 3 and 4

Behavioural results

To compare behavioural accuracy with the normally-hearing children tested in Experiment 3, a 2 x 2 mixed ANOVA was conducted with the factors hearing group (normally-hearing/hearing-impaired) and cue type (Location/Gender), separately for the Test and Control Conditions. In the Test Condition, normally-hearing children achieved significantly higher accuracy than hearing-impaired children [Figure 6.14; $F(1, 35) = 51.02$, $p < 0.001$, $\eta_p^2 = 0.59$]. However, there was no significant main effect of cue type [$F(1, 35) = 2.41$, $p = 0.13$] and no significant interaction between hearing group and cue type [$F(1, 35) = 2.05$, $p = 0.16$]. In the Control Condition, there was no significant difference in accuracy between normally-hearing (M = 94.7%, SD = 7.1) and hearing-impaired children (M = 98.1%, SD = 3.9), $F(1, 35) = 3.54$, $p = 0.07$. There was no significant main effect of cue type [$F(1, 35) = 3.09$, $p = 0.09$] and no significant interaction [$F(1, 35) = 1.44$, $p = 0.24$].

Event-related potentials

Figure 6.16 displays the average ERPs (averaged across Location and Gender Conditions and across broad posterior and anterior channel groups) that occurred in the Test and Control Conditions, separately for hearing-impaired children and the normally-hearing children who participated in Experiment 3. For each waveform, amplitudes at each time point were compared between normally-hearing and hearing-impaired children in an independent-samples *t*-test. When a Bonferroni correction was applied for multiple comparisons at 4700 time points, none of the time points reached the $p < 0.05$ criterion. This implies that there were no differences between the average waveforms for normally-hearing and hearing-impaired children. However, the Bonferroni correction might have been overly stringent, particularly with the small sample size in the hearing-impaired group. In order to estimate where differences between normally-hearing and hearing-impaired children might possibly occur, a $p < 0.01$ criterion was applied to the uncorrected *p*-values. The uncorrected *p*-values were intended to be informative, but did not form the basis for subsequent conclusions.

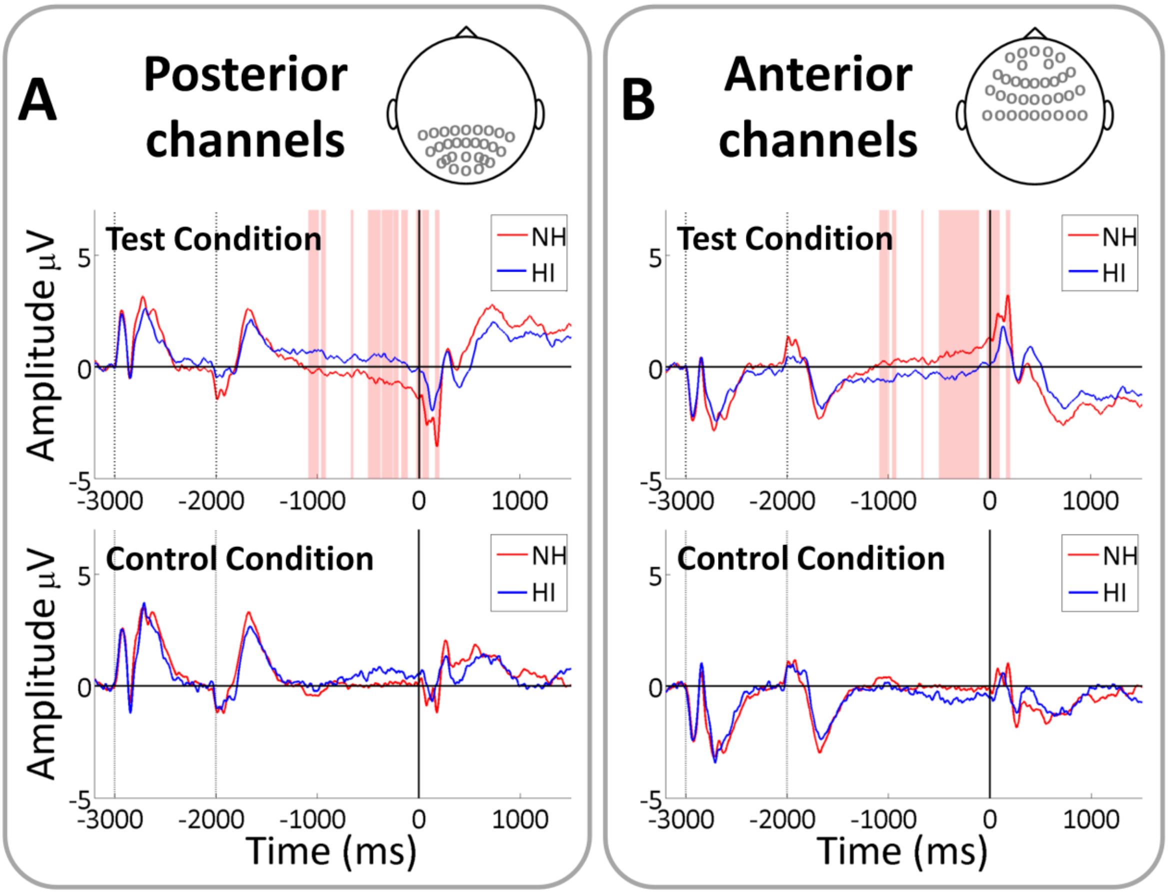


Figure 6.16. Experiments 3 and 4: Overall ERP waveforms for hearing-impaired children (HI) performing the task without their hearing aids and normally-hearing children (NH) who participated in Experiment 3. The waveforms have been averaged across cue types (Location/Gender), separately for the Test and Control Conditions. Panel (**A**) displays amplitudes averaged across a group of posterior channels. Panel (**B**) displays amplitudes averaged across a group of anterior channels. The electrodes included in each average are displayed on the scalp maps at the top of each panel. For each graph, the red shaded boxes indicate time points in which an independent-samples *t*-test revealed an uncorrected *p*-value that reached the 0.01 uncorrected criterion.

There were no significant differences between normally-hearing and hearing-impaired children at posterior or anterior channels during the Control Condition ($p > 0.01$, uncorrected). During the Test Condition, there were no significant differences during thebaseline period, during the presentation of the visual composite stimulus, or during the initial response to the reveal of the visual cue ($p > 0.01$, uncorrected). However, differences between normally-hearing and hearing-impaired children occurred during the second half of the Preparatory Phase (between -1092 and -923 ms, between -676 and -658 ms, between -505 and -109 ms, and between -33 and 0 ms). Differences also occurred at the beginning of

Table 6.4. Summary of comparisons between the Type-I clusters identified from normally-hearing children in Experiment 3 and amplitudes—averaged over the electrodes and time points that contribute to the cluster—in hearing-impaired children in Experiment 4. A tick in the column headed 'Significant difference between Test/Control Conditions?' indicates that a paired-samples *t*-test revealed a significant difference in amplitude between the Test and Control Conditions in the hearing-impaired children who participated in Experiment 4. A tick in the column headed 'Significant interaction between hearing groups?' indicates that a 2 x 2 between-subjects ANOVA with the factors hearing group (normally-hearing/hearing-impaired) and condition (Test/Control) revealed a significant two-way interaction (*p*-values displayed underneath).

Cue Type	Phase	Cluster Number	Significant difference between Test/Control Conditions?	Significant interaction between hearing groups?
Location	Preparatory	**1**	✗ $p = 0.66$	✗ $p = 0.15$
		2	✗ $p = 0.69$	✓ $p = 0.013$
		3	✗ $p = 0.64$	✓ $p = 0.026$
	Selective	**4**	✓ $p = 0.040$	✗ $p = 0.32$
		5	✗ $p = 0.09$	✗ $p = 0.49$
		6	✗ $p = 0.17$	✗ $p = 0.71$
Gender	Preparatory	**7**	✗ $p = 0.38$	✗ $p = 0.49$
	Selective	**8**	✓ $p < 0.001$	✗ $p = 0.53$
		9	✓ $p = 0.001$	✗ $p = 0.54$
		10	✗ $p = 0.33$	✗ $p = 0.06$
		11	✗ $p = 0.39$	✓ $p = 0.046$

the Selective Phase, during the initial responses to the onset of acoustical stimuli (between 0 and 95 ms and between 157 and 200 ms), but did not occur later during the Selective Phase. The second ERP analyses compared directly the average amplitude of ERPs for each of the clusters identified in normally-hearing children (reported in Experiment 3) between hearing-impaired children and the normally-hearing children who participated in Experiment 3. This analysis aimed to investigate whether hearing-impaired children showed atypical EEG activity compared to normally-hearing children. Figure 6.17 visualises the average

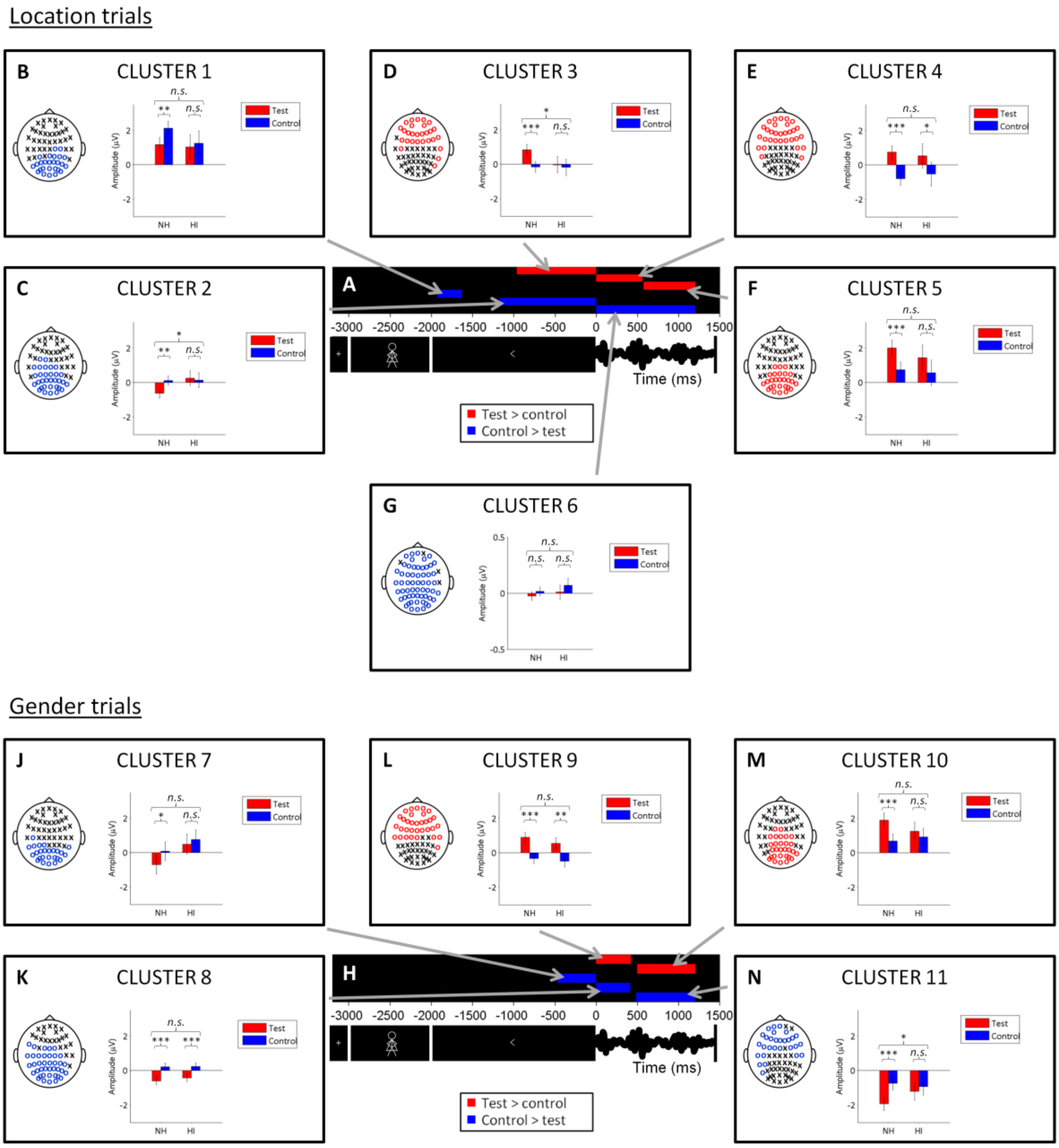


Figure 6.17. Experiments 3 and 4: Comparison of amplitudes in each cluster identified in the Type-I Spatio-temporal Cluster-based Permutation Analyses of Experiment 3 for Location (**A** to **G**) and Gender (**H** to **N**) trials between normally-hearing and hearing-impaired children when performing the task without their hearing aids. (**A** and **H**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. Further information about each cluster is displayed in (**B**)-(**G**) and (**J**)-(**N**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the bar graph shows the average amplitude of ERPs—averaged over the electrodes and time points that contributed to the cluster—for the normally-hearing (NH) children in Experiment 3 and the hearing-impaired children (HI) in Experiment 4 (correct-and-incorrect trials). Error bars show within-subjects 95% confidence intervals. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing the Test and Control Conditions within each hearing group (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). The larger brackets at the top of each graph indicate the significance level of the interaction between hearing group and the Test/Control Conditions.

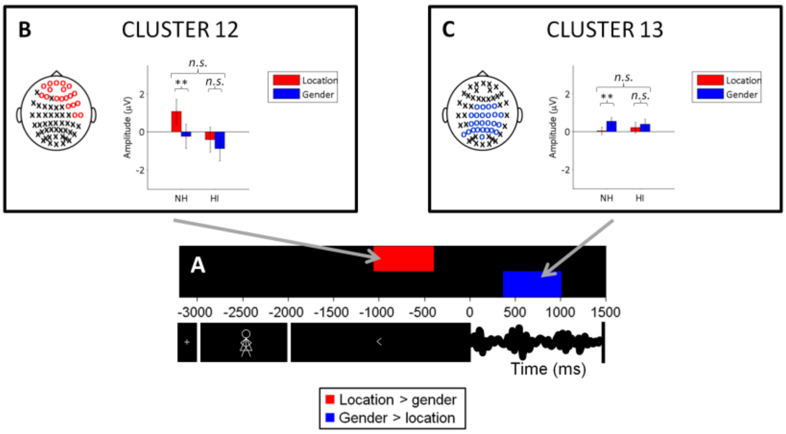


Figure 6.18. Experiments 3 and 4: Comparison of amplitudes in each cluster identified in the Type-II Spatio-temporal Cluster-based Permutation Analyses of Experiment 3, which contrasted Location and Gender trials in the Test Condition, between normally-hearing and hearing-impaired children. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. Further information about each cluster is displayed in (**B**)-(**C**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the bar graph shows the average amplitude of ERPs—averaged over the electrodes and time points that contributed to the cluster—for the normally-hearing (NH) children in Experiment 3 and hearing-impaired children (HI) in Experiment 4 who completed the task without their hearing aids (correct-and-incorrect trials). Error bars show within-subjects 95% confidence intervals. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing the Test and Control Conditions within each hearing group (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). The larger brackets at the top of each graph indicate the significance level of the interaction between hearing group and cue type (Location/Gender).

Table 6.5. Summary of comparisons between the Type-II clusters identified from normally-hearing children in Experiment 3 and amplitudes—averaged over the electrodes and time points that contribute to the cluster—in hearing-impaired children in Experiment 4. A tick in the column headed 'Significant difference between Location/Gender trials?' indicates that a paired-samples *t*-test revealed a significant difference in amplitude between Location and Gender trials in the hearing-impaired children who participated in Experiment 4. A tick in the column headed 'Significant interaction between hearing groups?' indicates that a 2 x 2 between-subjects ANOVA with the factors hearing group (normally-hearing/hearing-impaired) and cue type (Location/Gender) revealed a significant two-way interaction (*p*-values displayed underneath).

Condition	**Phase**	**Cluster Number**	**Significant difference between Location/Gender trials?**	**Significant interaction between hearing groups?**
Test	Preparatory	**12**	✗ p = 0.29	✗ p = 0.23
	Selective	**13**	✗ p = 0.36	✗ p = 0.16

amplitudes for Type-I clusters in normally-hearing and hearing-impaired participants. First, the average amplitudes of ERPs for each cluster were compared between the Test and Control Conditions in hearing-impaired children. The p-values resulting from paired-samples t-tests are tabulated in the fourth column of Table 6.4. The fifth column of Table 6.4 tabulates the p-values that resulted from a two-way interaction between hearing group (normally-hearing/hearing-impaired) and condition (Test/Control) in a 2 x 2 between-subjects ANOVA. Together, the results from the paired-samples t-test in hearing-impaired children and the results of the two-way interaction across hearing groups were used to inform conclusions about whether the patterns of ERPs were similar or different between normally-hearing and hearing-impaired children (Figure 6.12).

For Location trials, there was weak evidence for atypical ERPs in hearing-impaired children early during the Preparatory Phase (Cluster 1) but strong evidence for atypical ERPs later during the Preparatory Phase (Clusters 2 and 3). For Gender trials, there was weak evidence for atypical ERPs towards the end of the Preparatory Phase where the significant cluster occurred for normally-hearing children (Cluster 7).

Early during the Selective Phase, there was strong evidence for similar patterns of ERPs in normally-hearing and hearing-impaired children in Location (Cluster 4) and Gender (Clusters 8 and 9) trials. The remainder of the clusters during the Selective Phase showed weak evidence for atypical ERPs (Clusters 5, 6, and 10), apart from one cluster (Cluster 11) that occurred towards the end of the Selective Phase in Gender trials, which showed strong evidence for atypical ERPs.

Equivalent comparisons for the Type-II clusters are visualised in Figure 6.18 and tabulated in Table 6.5. The clusters during the Preparatory and Selective Phases both showed weak evidence for atypical ERPs between Location and Gender trials in hearing-impaired children.

6.6.3. Discussion

Hearing-impaired children in the current experiment showed fewer significant clusters of EEG activity than the normally-hearing children who participated in Experiment 3. During the Preparatory Phase, the Cluster-based Permutation Analysis revealed one significant cluster that differed between the Test and Control Conditions early during Gender trials (this cluster began at the same time that the visual cue was fully revealed; Cluster 14; Figure 6.15B) and no significant clusters during Location trials (Figure 6.15C). During the Selective Phase, there was one cluster that differed significantly between the Test and Control Conditions in the Location Condition and two in the Gender Condition (Figure 6.15).

Table 6.6. Conclusions from the comparisons between normally-hearing and hearing-impaired children for the Type-I and Type-II clusters identified in normally-hearing children in Experiment 3. Conclusions are based on the results reported in Tables 6.4 and 6.5 and the hypothetical patterns of results visualised in Figure 6.12.

Cue Type	Phase	Cluster Number	Conclusion
Location	Preparatory	**1**	**Weak evidence for atypical ERPs**
		2	**Strong evidence for atypical ERPs**
		3	**Strong evidence for atypical ERPs**
	Selective	**4**	**Strong evidence for similar ERPs**
		5	**Weak evidence for atypical ERPs**
		6	**Weak evidence for atypical ERPs**
Gender	Preparatory	**7**	**Weak evidence for atypical ERPs**
	Selective	**8**	**Strong evidence for similar ERPs**
		9	**Strong evidence for similar ERPs**
		10	**Weak evidence for atypical ERPs**
		11	**Strong evidence for atypical ERPs**
Test	Preparatory	**12**	**Weak evidence for atypical ERPs**
	Selective	**13**	**Weak evidence for atypical ERPs**

In addition to differences in the number of clusters that emerged from hearing-impaired children and normally-hearing children, some aspects of EEG activity showed significantly different patterns of activity between the two groups (Figure 6.17 and Table 6.4). The conclusions drawn from this evidence are tabulated in Table 6.6. There was evidence for atypical activity in hearing-impaired children throughout the Preparatory Phase. The strongest evidence for atypical activity arose towards the end of the Preparatory Phase in Location trials, immediately before the talkers started speaking. This finding is consistent with the idea that fewer significant clusters for hearing-impaired than normally-hearing children reflected atypical brain activity, rather than lower statistical power for the Cluster-based Permutation Analysis as a result of fewer hearing-impaired than normally-hearing participants.

Comparisons during the Selective Phase showed similarities in aspects of the response to the acoustical stimuli in normally-hearing and hearing-impaired children (Table 6.6). The clusters that occurred early (< 600 ms) after the talkers started speaking showed strong evidence for similar patterns of ERPs (Clusters 4, 8, and 9). This finding is likely to reflect similar responses to broad aspects of the acoustical stimuli that differed between the

Test and Control Conditions—for example the presentation of intelligible speech in the Test Condition, but noise-vocoded stimuli in the Control Condition, or acoustical stimuli presented from three different spatial locations in the Test Condition compared to one location in the Control Condition. This result provides evidence that the stimuli were audible for hearing-impaired children, so differences between normally-hearing and hearing-impaired children cannot be attributed to a lack of audibility for hearing-impaired children.

Later during the Selective Phase, when EEG activity was likely to reflect selective attention to key words spoken by one talker, there was strong evidence for atypical ERPs in

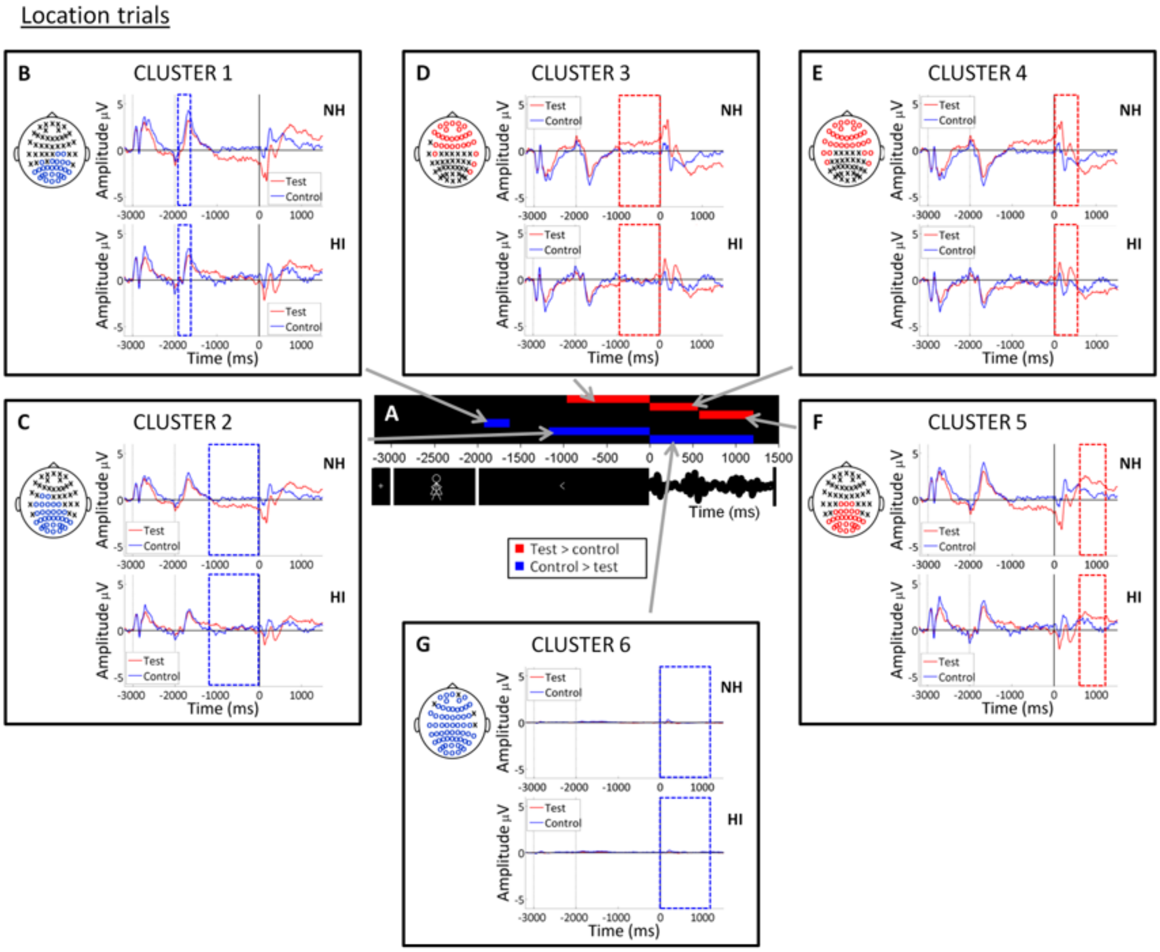


Figure 6.19. (Continued on next page). Experiments 3 and 4: Comparison of ERP waveforms at the electrodes that contributes to each cluster identified in the Type-I spatio-temporal cluster-based permutation analyses of Experiment 3 for the Location Condition (**A** to **G**) and the Gender Condition (**J** to **N**). (**A** and **H**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. ERP waveforms—averaged across the electrodes that contributed to each cluster—are displayed separately for normally-hearing (NH) and hearing-impaired (HI) children in (**B**)-(**G**) and (**J**)-(**N**). For each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the time-span of the cluster is indicated on the ERP waveforms by a dashed rectangle.

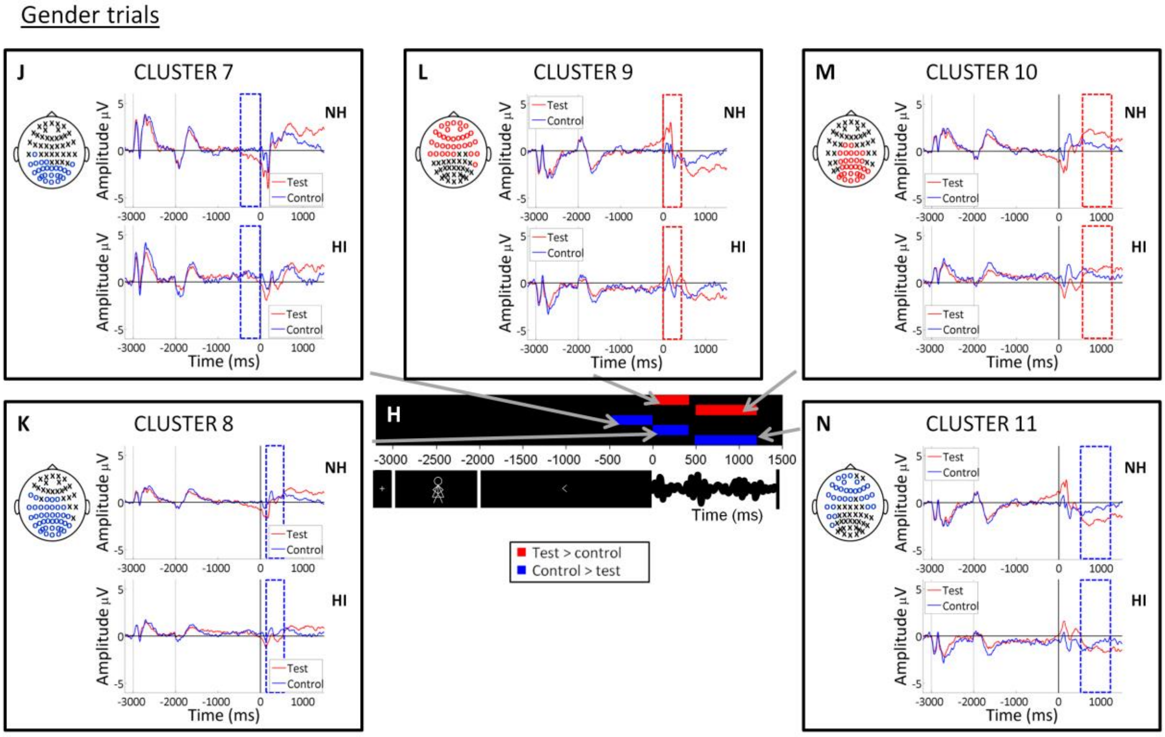


Figure 6.19. (Continued from previous page).

Gender trials and weak evidence for atypical ERPs in Location trials. These differences could *either* reflect differences in selective attention between the groups *or* differences in responses to the acoustical stimuli due to differences in peripheral transduction.

Although the current experiment only statistically analysed differences in the *amplitudes* of ERPs between normally-hearing and hearing-impaired children, changes in the latencies of ERPs do not seem to underlie the results reported. Figure 6.16 shows similarities in the latencies at which the overall ERP waveforms reached peak amplitude in normally-hearing and hearing-impaired children (when ERPs were examined in broad groups of electrodes). In addition, Figures 6.19 and 6.20 show that the waveforms within each of the clusters identified in normally-hearing children had similar latencies in normally-hearing and hearing-impaired children. Consequently, for the results reported in this experiment, differences between normally-hearing and hearing-impaired children resulted from differences in the amplitudes of responses, rather than their associated latencies.

6.7. General discussion

The experiments reported in this chapter had two aims. First, to measure the extent to which the accuracy and latency of reporting key words spoken by a target talker were affected by the duration of preparation time in normally-hearing and hearing-impaired children. Second, to measure brain activity using EEG during three-talker listening in normally-hearing and hearing-impaired children.

Overall, Experiment 1 provides evidence that normally-hearing children achieve higher accuracy of speech intelligibility when they have time to prepare for the location or gender of an upcoming talker compared to when they have no time to prepare before the talkers start speaking (Figure 6.4A). Experiment 2 showed that hearing-impaired children did *not* achieve higher accuracy of speech intelligibility when they had time to prepare for an upcoming talker, although the latency of speech intelligibility was progressively shorter as the duration of preparation time increased (Figure 6.6).

Experiment 3 revealed significant EEG activity during three-talker listening in

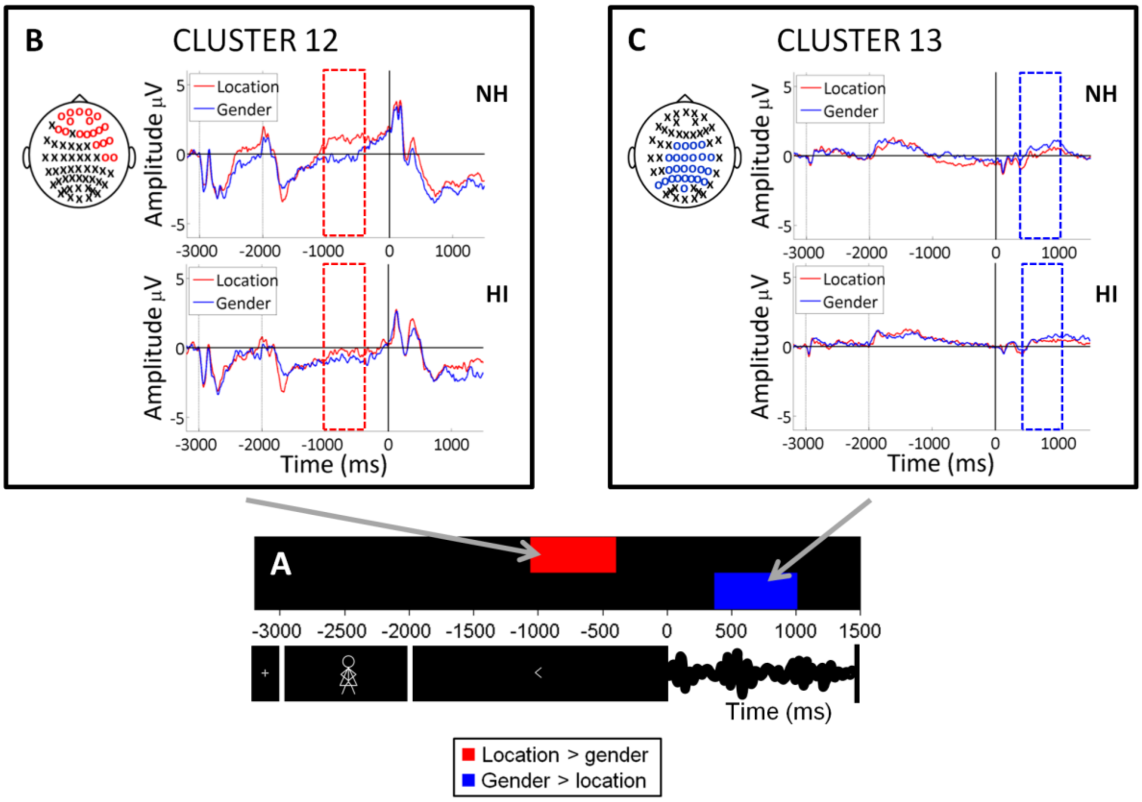


Figure 6.20. Experiments 3 and 4: Comparison of ERP waveforms at the electrodes that contributes to each cluster identified in the Type-II spatio-temporal cluster-based permutation analyses of Experiment 3. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. Time on the x-axis is relative to the onset of the acoustical stimuli. ERP waveforms—averaged across the electrodes that contributed to each cluster—are displayed separately for normally-hearing (NH) and hearing-impaired (HI) children in (**B**)-(**C**). For each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the time-span of the cluster is indicated on the ERP waveforms by a dashed rectangle.

normally-hearing children. This activity occurred throughout the Preparatory and Selective Phases when participants were cued either to the location or to the gender of an upcoming talker (Figure 6.9). Experiment 4 identified aspects of EEG activity that were atypical in hearing-impaired children and aspects of EEG activity that were similar to activity observed in normally-hearing children. Strong evidence for *atypical* EEG activity occurred during the second half of the Preparatory Phase, while strong evidence for *similar* EEG activity occurred early during the Selective Phase, at the same time as the talkers started speaking (Table 6.6).

6.7.1. Comparisons between normally-hearing and hearing-impaired children

As expected, hearing-impaired children showed lower accuracy of speech intelligibility than normally-hearing children. This result is consistent with previous experiments that have found that hearing-impaired listeners struggle in noisy environments (e.g. Dubno et al., 1984; Salvi et al., 2002) and achieve lower speech intelligibility during multi-talker listening than normally-hearing listeners (Helfer & Freyman, 2008; Marrone et al., 2008a). The finding of lower accuracy also suggests that these experiments accessed an aspect of speech intelligibility that hearing-impaired children struggle with. Nevertheless, the hearing-impaired children still performed significantly above chance in both experiments (Experiments 2 and 4), which suggests that they were not simply guessing and had not given up on the task.

6.7.1.1. How preparation time affects the accuracy and latency of speech intelligibility

In contrast to normally-hearing children, hearing-impaired children did not show improved accuracy of speech intelligibility when they had time to prepare for an upcoming talker (Figure 6.6). This finding is consistent with the results of Best *et al.* (2009), who presented a visual cue to indicate the location of an upcoming talker during multi-talker listening. They compared the accuracy of speech intelligibility for adults with normal hearing and adults with moderate hearing loss between trials in which they received a cue for location and trials in which they received no cue. Similar to the current experiment, they found that hearing-impaired listeners did not receive as much improvement in the accuracy of speech intelligibility as normally-hearing listeners when they received a visual cue for location.

6.7.1.2. Ability to segregate competing talkers

When normally-hearing and hearing-impaired children did not correctly identify words spoken by the target talker, the majority of errors consisted of words spoken by one of the competing talkers—either the opposite-gender talker or the child talker. This result suggests that the main source of difficulty for *both* normally-hearing and hearing-impaired children was selecting the correct talker to which to attend (i.e. difficulties in 'object selection'), rather than difficulties segregating the talkers (i.e. difficulties in 'object formation').

However, there was a significant difference in the overall distribution of error types between hearing-impaired and normally-hearing children. Hearing-impaired children made a higher percentage of errors than normally-hearing children that consisted of words spoken by a mixture of talkers. This finding shows that, although the main cause of difficulty was directing selective attention to a mixture of talkers (i.e. difficulties in 'object selection'), hearing-impaired children were significantly more likely than normally-hearing children to have difficulties segregating the talkers or grouping words spoken by the same talker (i.e. difficulties in 'object formation'). Failures to segregate simultaneous talkers might arise when the spectro-temporal features of the target are not easily distinguishable from features of the maskers, which is more likely to occur in hearing-impaired listeners than normally-hearing listeners due to poorer frequency selectivity (e.g. Festen & Plomp, 1983) and deficits extracting or encoding temporal fine structure (Lorenzi et al., 2006). Hearing-impaired children were also more likely to report words that were not spoken on the current trial than normally-hearing children, which suggests higher energetic making of the target talker in hearing-impaired than normally-hearing children. Overall, these findings indicate multiple causes of possible difficulty for speech intelligibility during multi-talker listening in hearing-impaired children.

Both normally-hearing and hearing-impaired children showed a similar overall pattern of errors across the four different cue types (left/right/male/female). However, there were subtle differences in the *balance* of errors between different cue types for both hearing groups. For example, on attend-male trials, participants were more likely to report words spoken by the female talker than the child talker (Figures 6.5 and 6.7). This finding is consistent with the idea that the fundamental frequencies (F0s) of the male and female talkers were more similar than those of the male and child talkers. Similarly, on attend-female trials, participants were most likely to report words spoken by the talker with the most similar F0, although the talker with the most similar F0 to the female talker was the child talker.

6.7.1.3. EEG activity during multi-talker listening

Overall ERP waveforms

During the Control Condition, hearing-impaired children showed similar patterns of average ERP waveforms as normally-hearing children (Figure 6.16). This finding demonstrates that there were no differences in EEG activity between the groups on a task in which participants had to respond to a visual stimulus.

Even though different criteria were adopted for analysing trials in normally-hearing and hearing-impaired children (incorrect trials were included in the analyses of ERPs for hearing-impaired children but not for normally-hearing children) and no correction for multiple comparisons was performed, the initial response to the reveal of the visual cue during the Test Condition was similar in normally-hearing and hearing-impaired children. Where there were differences in activity using the uncorrected criterion, those differences occurred later during the Preparatory Phase and early during the Selective Phase (Figure 6.16). During parts of the waveform that reached the uncorrected criterion, the amplitudes for hearing-impaired children were generally closer to the baseline amplitude than for normally-hearing children. In general, the finding of lower-amplitude ERPs for hearing-impaired than normally-hearing children during the presentation of acoustical stimuli is consistent with the results of previous experiments. For example, Koravand, Jutras, and Lassonde (2012) measured evoked potentials for pure tones and speech syllables in children with normal hearing and children with mild and moderate hearing loss, all aged 9–10 years. They found that the amplitude of the N2 component was significantly reduced in hearing-impaired compared to normally-hearing children. This result is consistent with the explanation that higher thresholds for detecting acoustical stimuli are associated with reduced amplitudes of ERPs.

Activity during the Preparatory Phase of Location and Gender trials

Normally-hearing children showed significant preparatory EEG activity during the second half of the Preparatory Phase, but hearing-impaired children did not (Figure 6.15). In addition, the clusters that occurred in normally-hearing children during the second half of the Preparatory Phase of Location trials (Clusters 2 and 3) showed significantly different amplitudes in normally-hearing and hearing-impaired children (Table 6.4). This finding suggests that lower-amplitude EEG activity in hearing-impaired children reflects atypical brain activity, rather than lower statistical power for detecting significant differences due to a small sample of hearing-impaired children. This atypical activity is likely to reflect difficulties preparing attention for an upcoming talker based on cues for location or gender in hearing-impaired children.

Activity during the Selective Phase of Location and Gender trials

During the Selective Phase, hearing-impaired children showed significant differences in ERPs between the Test and Control Conditions (Figure 6.15). This result suggests that broad differences in the acoustical stimuli between the Test and Control Conditions were reflected in the amplitudes of ERPs in hearing-impaired children, even when they performed the task without their hearing aids. In addition, the activity that occurred early during the Selective Phase was similar in hearing-impaired children to the activity shown by normally-hearing children (Table 6.4).

However, later during the Selective Phase, there was evidence for atypical ERPs in hearing-impaired children (Table 6.6). There are at least two possible explanations for this finding. First, atypical peripheral transduction in hearing-impaired children could affect the brain activity evoked by differences in the acoustical stimuli between the Test and Control Conditions. Second, neural activity that underlies attentional selection may differ between normally-hearing and hearing-impaired children. The current experiment is not able to distinguish these alternatives. However, both factors have the potential to contribute to poorer speech intelligibility in hearing-impaired than normally-hearing children.

6.7.2. Possible limitations

6.7.2.1. Differences in level between normally-hearing and hearing-impaired children

The procedure was intended to be identical for normally-hearing and hearing-impaired children. One difference, however, was a higher presentation level for hearing-impaired than normally-hearing children. The reason for this difference was that the presentation level used for normally-hearing children made the stimuli inaudible for a hearing-impaired child who was tested at this level in a pilot experiment. These experiments aimed to identify patterns of accuracy and ERPs that were not due to differences in stimulus audibility and, therefore, the stimuli were presented at a higher level for hearing-impaired children, for which the stimuli were expected to be audible for participants with moderate hearing loss. The difference in stimulus presentation level, however, is unlikely to have affected the results reported in this chapter. For example, there is no reason to believe that accuracy and RTs in normally-hearing children would have differed between average stimulus presentation levels of 63 and 76 dB: first, because the target and competing talkers were always presented at equal levels and second, because both levels would be audible for normally-hearing participants. In addition, differences in ERPs between normally-hearing and hearing-impaired children showed smaller between-condition amplitude differences in hearing-impaired than normally-hearing children; the opposite pattern would be expected if differences in presentation levels were reflected in differences in ERPs between the groups (e.g. Hegerl, Gallinat, & Mrowinski, 1994; Rapin, Schimmel, Tourk, Krasnegor, & Pollak, 1966; Schadow et al., 2007).

6.7.2.2. Small sample size of hearing-impaired children

The experiments reported in this chapter aimed to recruit the same number of hearing-impaired as normally-hearing children. However, due to difficulties recruiting patients who fit the criteria for participation (bilateral moderate cochlear hearing loss; aged between 7 and 16 years; native English speakers; and no physical or cognitive disabilities that would prevent understanding of, or participation in, the listening tasks), Experiments 2 and 4 report data from fewer hearing-impaired children than Experiments 1 and 3 report for normally-hearing children. The small sample size had two main consequences. First, the analyses of differences between conditions *within* the hearing-impaired group might have been underpowered—particularly for the EEG results reported in Experiment 4 (Figure 6.15). Second, comparisons between normally-hearing and hearing-impaired children may have been underpowered—particularly for the interactions between normally-hearing and hearing-impaired children reported in Experiments 2 (Section 6.4.2.2) and 4 (Tables 6.4 and 6.5). Consequently, the results reported may underestimate the extent of differences in speech intelligibility and differences in EEG activity between normally-hearing and hearing-impaired participants. Nevertheless, the current experiments were still able to identify significant differences between conditions within hearing-impaired children and significant interactions between normally-hearing and hearing-impaired children. These significant results are likely to reflect large effects that were consistent across the participants from which data were collected. Additional participants are required to confirm that the observed effects persist across larger numbers of hearing-impaired children.

Nevertheless, of the small sample of hearing-impaired participants and wide age range of children tested, the patterns of results were similar across participants. Appendix B reports the results from individual participants in Experiments 1–4. Although participants achieved higher accuracy of speech intelligibility with increasing age, the patterns for the accuracy and latency of speech intelligibility with increasing preparation time were relatively constant across participants (Experiments 1 and 2). In addition, age did not correlate significantly with the amplitude of significant clusters of EEG activity identified during the Preparatory or Selective Phases (Experiments 3 and 4). Overall, these results suggest that the results reported in this chapter were consistent across individual participants of different ages.

6.7.2.3. EEG recordings without hearing aids

The results reported for Experiments 2 and 4 were gathered when participants with hearing loss completed the tasks *without their hearing aids*. Since all of the children listened with their hearing aids in everyday life more often than they listened without their hearing aids, one possibility was that differences between normally-hearing and hearing-impaired children could be explained by unfamiliar listening conditions for hearing-impaired children. In addition, it was possible that, even though the acoustical stimuli were presented at higher levels for hearing-impaired than normally-hearing children, the acoustical stimuli were still inaudible for children with the highest audiometric thresholds. To rule out these possible explanations, hearing-impaired participants were tested in a second session of Experiment 4 in which they completed an identical task using their own acoustic hearing aids. The results are reported in Appendix C. The results showed largely similar patterns of ERPs when hearing-impaired children completed the task with and without their hearing aids. Therefore, inaudibility or unfamiliar listening conditions are unlikely to fully explain the differences between normally-hearing and hearing-impaired children reported in this chapter.

6.7.2.4. Inclusion of incorrect trials

Another consideration when interpreting the results of Experiment 4 is the inclusion of correct-only or correct-and-incorrect trials in the analyses. Ideally, the analyses for normally-hearing and hearing-impaired children would have both only included correct trials—because differences in ERPs gained from correct-and-incorrect trials could reflect differences in behavioural performance, rather than the EEG activity that accompanied successful trials (which has the potential to produce confounds, for example, if one group was not engaged in the task for all trials of the experiment; described in more detail in Appendix D). However, the statistical power for detecting differences between normally-hearing and hearing-impaired children was lower when correct-only trials were analysed due to fewer trials. To ensure that the reported effects could not be explained by the inclusion of incorrect trials in the analyses for hearing-impaired children, the analyses reported in Appendix D compared the results from correct-and-incorrect and correct-only trials. The results provided evidence for similar patterns of amplitudes for correct-and-incorrect and correct-only trials in hearing-impaired children. This finding suggests that differences between normally-hearing and hearing-impaired children reported in the current chapter cannot be explained by the inclusion of incorrect trials in the analyses for hearing-impaired children.

6.7.2.5. Analysis for comparing ERPs between normally-hearing and hearing-impaired children

It is worth noting that the analyses that compared ERPs directly between normally-hearing and hearing-impaired children (Section 6.6.2.3) analysed clusters that were selected from the data from normally-hearing children. Importantly, the conclusion of atypical ERPs in hearing-impaired children were based on the results of two different analyses, the first being the finding that the Spatio-temporal Cluster-based Permutation Analysis (Section 6.6.2.2) for hearing-impaired children revealed only one significant cluster during the Preparatory Phase for the Type-I analysis (Figure 6.15). The second piece of evidence (Section 6.6.2.3) aimed to test whether the clusters identified in normally-hearing children were also present in hearing-impaired children, but did not emerge from the analysis reported in Section 6.6.2.2 due to lower power for detecting significant clusters in hearing-impaired children. One possible argument is that the latter analysis was biased because the clusters for comparison were identified and subsequently tested in the normally-hearing children. Although, alternative analyses may have been able to have been employed if a greater number of children were tested.

If a greater number of normally-hearing children were tested in Experiment 3, then it would have been possible to split the data from normally-hearing children into two groups, each containing half of the total number of participants. The data from the first group could be analysed using the Cluster-based Permutation Analysis. Subsequently, amplitudes from the electrodes and time points that contributed to significant clusters could be compared directly between the second group of normally-hearing children and the hearing-impaired children. This method would improve upon the method reported in Section 6.6.2.3, although with the current number of participants in Experiment 3, there would not be sufficient power to identify significant clusters if the group were to be split in half.

A second possible method would be to combine the data from an equal number of normally-hearing and hearing-impaired children and perform the Cluster-based Permutation Analysis on the combined data. Subsequently, amplitudes could be tested on those clusters between the normally-hearing and hearing-impaired children. Although, a possible limitation of that approach is that significant activity that occurred in the normally-hearing but not in the hearing-impaired children might be cancelled out by combining the two groups for the Cluster-based Permutation Analysis. Consequently, the results would likely underestimate the extent of differences between normally-hearing and hearing-impaired children.

6.7.3. Conclusions

During three-talker listening, normally-hearing children achieved improved accuracy of speech intelligibility when they had time to prepare for the location or gender of an upcoming talker than when they had no time to prepare. Consistent with this finding, normally-hearing children displayed significant preparatory EEG activity when they were cued either to the location or to the gender of a target talker during three-talker listening. The results showed extensive similarities in the activity evoked when participants were cued to location or gender, which provides evidence for domain-general preparatory activity. Although, there were also some aspects of EEG activity that were significantly different when participants attended to location or gender. Overall, the results demonstrate that normally-hearing children aged 7–16 years display similar domain-general and cue-specific EEG activity to adults when they are cued to the location or gender of a target talker during three-talker listening.

Overall, ERP waveforms were similar in hearing-impaired and normally-hearing children. However, there was strong evidence for atypical EEG activity for hearing-impaired children immediately before the talkers started speaking, which reflected an absence of preparatory activity in hearing-impaired children. This finding is consistent with the idea that hearing-impaired children do not utilise preparatory attention to the same extent as normally-hearing children. Therefore, the results suggest that atypical preparatory attention might be one factor that contributes to poorer speech intelligibility in hearing-impaired children during multi-talker listening.

Chapter 7
Summary and General Discussion

7.1. Recap of research aims

This thesis examines preparatory and selective attention during multi-talker listening by participants with normal and impaired hearing. In more detail, preparatory and selective attention refer to the mechanisms by which participants prepare their attention when they know attributes of a talker before he or she begins to speak and the mechanisms by which participants attend selectively to a talker while multiple talkers speak simultaneously. Chapter 1 reviews previous research and defines the rationale for the aims of this thesis. Chapter 2 describes the method of electro-encephalography (EEG) and the analysis techniques used throughout this thesis.

The experiments reported in Chapters 3–5 had two main aims: (1) to devise a technique for measuring preparatory and selective attention during multi-talker listening in normally-hearing adults, which would also be suitable for normally-hearing and hearing-impaired children, and (2) to examine the time course of preparatory and selective attention in normally-hearing listeners. To this aim, normally-hearing adults were tested on two different multi-talker listening tasks—one in which *two* talkers spoke simultaneously and the other in which *three* talkers spoke simultaneously. In these experiments, the task was to report words spoken by a target talker who was specified by either their location (left/right) or their gender (male/female). The experiments reported in Chapter 4 investigated how the duration of preparation time affected the accuracy and latency of speech intelligibility; Chapters 3 and 5 measured EEG activity evoked in these tasks. Of key interest were the timing of EEG activity during multi-talker listening and aspects of the response that showed domain-general or cue-specific attention when participants were cued to the location or gender of a target talker.

The experiments reported in Chapter 6 aimed to investigate whether children with moderate bilateral cochlear hearing loss show atypical preparatory attention. Two of the experiments reported in Chapter 6 investigated how the duration of preparation time affected the accuracy and latency of speech intelligibility in normally-hearing and hearing-impaired children. Two further experiments measured EEG activity evoked during three-talker listening in these groups of participants.

7.2. Summary of findings

7.2.1. Main findings of Chapter 3

- Before two talkers started speaking (i.e. during the 1000-ms Preparatory Phase), normally-hearing adults showed significant EEG activity when they were cued to the location (left/right) of the target talker. Preparatory EEG activity occurred early (< 50 ms) after a visual cue for location was revealed.
- Visual cues that indicated a target talker's gender led to significant preparatory EEG activity only when the cue predicted the *identity* of the target talker, but not when the cue specified only the gender of the talker without also predicting their identity. When significant activity occurred during the Preparatory Phase of gender trials, the timing and scalp distribution was similar to that found in location trials, which provides evidence for domain-general preparatory brain activity (i.e. activity that is similar when participants are cued to either location or gender).
- During the Preparatory Phase, there was *no* evidence for cue-specific activity that could be attributed to differences in the mechanisms that participants use to prepare their attention based on knowledge of a target talker's location compared to their gender. One possible explanation was that the two-talker task was sufficiently easy that participants did not need to deploy cue-specific attention during the Preparatory Phase in order to achieve accurate speech intelligibility.
- During the Selective Phase (i.e. while two talkers spoke simultaneously), the results showed consistent cue-specific activity that depended on whether participants attended selectively to the target talker on the basis of their location or their gender.
- Overall, patterns of ERPs recorded from children, aged 7–13 years, during the two-talker listening task were similar to adults. They showed significant EEG activity during the Preparatory Phase when they received a cue for location and

significant cue-specific activity during the Selective Phase depending on whether they attended selectively to a talker based on their location or their gender. However, one difference was that children did not display significant preparatory activity when they received a cue that indicated the *gender* of the target talker.

7.2.2. Main findings of Chapter 4

- For normally-hearing adults performing a two-talker listening task, accuracy of speech intelligibility was near ceiling even when participants had no time to prepare before the talkers began. Therefore, no significant improvements in the accuracy or latency of speech intelligibility were observed with increasing durations of preparation time. This finding was thought to underlie the results reported in Chapter 3, which found no evidence for significant cue-specific EEG activity during the Preparatory Phase when participants received cues for location compared to gender.
- However, in a more challenging three-talker task, longer preparation times (over the range of 0 to 2000 ms) progressively improved the accuracy and latency of speech intelligibility for normally-hearing adults. This finding is inconsistent with the idea of a 'threshold' of preparation time required for successful attentional preparation.

7.2.3. Main findings of Chapter 5

- During a three-talker listening task, normally-hearing adults showed significant EEG activity during the 2000-ms Preparatory Phase—when they received cues for both the location or for the gender of a target talker. Preparatory EEG activity occurred in two phases: (1) with an early latency (< 25 ms) after the visual cue was fully revealed, and (2) in the 750-ms interval immediately before the talkers began. Similarities in the EEG activity that occurred on location and gender trials provide evidence for domain-general preparatory attention.
- In addition to broadly similar activity, there was also cue-specific activity during some parts of the task, during which activity differed significantly between location and gender trials. These differences occurred during the middle of *both* the Preparatory and Selective Phases.

7.2.4. Main findings of Chapter 6

- Normally-hearing children, aged 7–16 years, showed improved accuracy of speech intelligibility when they had time to prepare for the location or gender of a target talker before three talkers began to speak. Like adults, normally-hearing children showed significant EEG activity during the Preparatory Phase that was similar when they were cued to the location or to the gender of a target talker. Similar to adults, there were also some significant differences in EEG activity between location and gender trials during the Preparatory and Selective Phases.

- Hearing-impaired children showed aspects of EEG activity that were similar and aspects of activity that were different to normally-hearing children. There were similar responses to the onset of acoustical stimuli (< 600 ms after the talkers started speaking). However, there was strong evidence for atypical EEG activity in hearing-impaired children during the second half of the Preparatory Phase, during which normally-hearing children displayed significant EEG activity but hearing-impaired children did not. It was inferred that this finding reflects atypical attentional preparation when hearing-impaired children receive cues that indicate the location or gender of a target talker.

7.3. General discussion

7.3.1. Domain-general and cue-specific EEG activity in normally-hearing adults

7.3.1.1. <u>Domain-general EEG activity</u>

Normally-hearing listeners showed consistent similarities in EEG activity when they were cued to the location compared to the gender of a target talker (Chapters 3 and 5), thus reflecting domain-general processing. The finding of domain-general preparatory activity when participants were cued to different attributes of an upcoming talker is consistent with the results of a functional magnetic resonance imaging (fMRI) experiment by Hill and Miller (2010). They cued participants to attend to either the location or to the fundamental frequency (F0) of an upcoming talker during three-talker listening. The results revealed activity in a highly-overlapping left-dominant fronto-parietal network when participants were cued to location or F0. The current results build upon the results of Hill and Miller by showing sustained domain-general attentional preparation throughout the Preparatory Phase (Chapters 3 and 5). This finding suggests that listeners utilise all of the time available to prepare their attention before a target talker begins to speak.

7.3.1.2. <u>Cue-specific EEG activity</u>

Although domain-general activity was observed throughout the task, there was also evidence for significant cue-specific EEG activity during some parts of the Preparatory and

Selective Phases (Chapter 5). Cue-specific activity during the Preparatory Phase was interpreted as evidence for differences in the processes by which participants prepare their attention for a target talker based on knowledge of the talker's location or gender. The experiment reported in Chapter 5 showed that this activity occurred approximately 1000 ms after the visual cue was revealed, in the middle of the Preparatory Phase. Overall, the finding of cue-specific preparatory EEG activity is consistent with previous experiments that have shown cue-specific brain activity when participants prepared for upcoming visual (e.g. Giesbrecht et al., 2006) and acoustical (e.g. Hill & Miller, 2010; Lee et al., 2013; Voisin et al., 2006) stimuli.

Cue-specific activity during the Selective Phase was found consistently across experiments and was interpreted as differences in the mechanisms that participants use to pick out a talker based on their location or gender while multiple talkers speak simultaneously. This activity occurred at approximately the time that the colour and number key words were spoken by the target talker (Chapters 3 and 5). Hill and Miller (2010) found cue-specific activity during the Selective Phase of their multi-talker listening task, although it was necessary for Hill and Miller to a select high-performing sub-set of their participants (for which accuracy was higher than 50% correct) in order to detect cue-specific activity. In the current experiments, normally-hearing adults achieved high (> 75%) accuracy, which might have been one factor that contributed to consistent observations of cue-specific activity during the Selective Phase across multiple experiments. Nevertheless, in some of the experiments reported in this thesis, behavioural accuracy was significantly higher for Location than Gender trials (for example, for Experiment 2 of Chapter 4). While every effort was made to equate accuracy in Location and Gender trials, a replication is desirable to confirm that cue-specific EEG activity did not result from small (but statistically significant) differences in accuracy.

Overall, the findings of both domain-general and cue-specific activity are consistent with previous experiments that have shown that spatial (location) and non-spatial (e.g. colour or F0) cues activate brain activity in an overlapping network of regions, although the detailed pattern of activity within this network depends on the specific attribute (spatial/non-spatial) to which participants attend. These findings have been shown for endogenous attention both to visual (e.g. Giesbrecht et al., 2006; Green & McDonald, 2008) and to acoustical non-speech stimuli (e.g. Ahveninen et al., 2006; Voisin et al., 2006).

7.3.2. Differences in ERPs between the two- and three-talker tasks

Compared to the two-talker experiment (Experiment 2, Chapter 3), the three-talker experiment (Chapter 5) revealed: (1) a greater number of significant clusters during the Preparatory Phase, (2) a longer duration of significant clusters, and (3) that significant clusters spanned a greater number of electrodes in the three-talker task compared to equivalent clusters in the two-talker task. The three-talker task was more difficult than the two-talker task and previous experiments have shown a greater magnitude of activity in the fronto-parietal network for difficult than easy tasks. For example, Falkenberg, Specht, and Westerhausen (2011) found greater magnitude of BOLD activity during trials in which participants were cued to report the less salient of two consonant-vowel stimuli during dichotic listening than when they were cued to the more salient stimulus. In theory, a greater magnitude of activity in the neural generators of EEG activity has the potential to manifest either as a greater number of significant scalp clusters or as significant differences that are sustained over a greater number of electrodes or time points. Therefore, the results of the current experiments are consistent with greater activity in the neural generators of the observed EEG activity in the three-talker than the two-talker task.

7.3.3. Comparisons between normally-hearing children and adults

These experiments aimed to provide a task that would be suitable for children. Behavioural accuracy was high for both adults and children, which suggests that the task was indeed appropriate for children.

In general, the results from normally-hearing children showed similar patterns to adults. For example, children showed improved accuracy and shorter latency of speech intelligibility when they had time to prepare for an upcoming talker (1000-ms and 2000-ms intervals) than when they had no time to prepare (0-ms interval; Experiment 1 of Chapter 6). This result is consistent with a previous experiment showing that children, like adults, can use advance cueing to improve the accuracy of speech intelligibility in noisy environments (Dhamani et al., 2013).

Consistent with the pattern of results for the accuracy and latency of speech intelligibility, there were also similarities in aspects of EEG activity between children and adults. Overall, the clusters that were identified in children had similar latencies and scalp distributions to the clusters identified in adults (for example, between Experiment 3 and Experiment 2 of Chapter 3; also, between Experiment 3 of Chapter 6 and the experiment reported in Chapter 5). These similarities suggest that children utilised aspects of brain activity during preparatory and selective attention that were similar to adults.

One difference was greater ERP amplitudes, on average, in children than adults, which has also been reported in previous experiments. Several possible reasons have been suggested, including thinner skulls in children than adults (McCullouch, 2013; Shapiro & Janzen, 1960) and smaller head sizes that increase the proximity of neural generators to recording electrodes (Picton & Taylor, 2007). For these reasons, greater overall amplitudes were not expected to relate specifically to differences in multi-talker listening between adults and children. Another consistent difference was that children generally showed fewer significant clusters than adults. However, it was unclear whether this finding arose from weaker preparatory and selective attention in children, since they typically identify speech in noise with lower accuracy than adults (Bonino et al., 2012; Fallon et al., 2000; Papso & Blood, 1989), or due to noisier ERPs in children with less statistical power for detecting significant differences. Nevertheless, overall, similarities between normally-hearing children and adults in these experiments were more extensive than the differences between them.

7.3.4. Preparatory attention in hearing-impaired children

Children with moderate bilateral hearing loss achieved lower accuracy of speech intelligibility than normally-hearing children in the same age range (Chapter 6). This result is consistent with previous experiments suggesting that hearing-impaired listeners struggle in noisy environments (e.g. Dubno et al., 1984; Salvi et al., 2002) and achieve lower speech intelligibility during multi-talker listening than normally-hearing listeners (Helfer & Freyman, 2008; Marrone et al., 2008a). Lower average accuracy is consistent with the idea that the three-talker listening task approximated aspects of everyday multi-talker environments that hearing-impaired listeners struggle with.

Hearing-impaired children did not show improved accuracy of speech intelligibility when they had time to prepare for an upcoming talker than when they did not (Experiment 2 of Chapter 6). This finding is consistent with the results of Best *et al.* (2009), who found that adults with moderate hearing loss received less benefit to the accuracy of speech intelligibility than normally-hearing adults when they were cued to the location of an upcoming talker. On each trial in Best *et al.*'s experiment, the target speech occurred in one of five time windows and at one of five spatial locations. Maskers, which consisted of time-reversed speech, were presented during the other time windows at the target's spatial location and at the other spatial locations during the target time window. In the 'where' cue condition, a light-emitting diode (LED) indicated the spatial location from which the target speech would occur. The intelligibility benefit of knowing the spatial location of the upcoming talker (calculated as the difference in percent correct between the 'where' cue condition and a no-cue condition) was significantly smaller for hearing-impaired listeners than for normally-hearing listeners. Therefore, like the results of the current experiments, the results of Best *et al.* suggest that hearing-impaired listeners did not benefit as much as normally-hearing listeners when they were able to prepare for the spatial location of an upcoming talker compared to when they did not know the location of the target talker.

The results from Chapter 6 extend the findings of Best *et al.* (2009) in four main ways: (1) they extend the results to children aged 9-16 with moderate hearing loss, (2) they demonstrate that hearing-impaired children also do not benefit from a cue for *gender* for the accuracy of speech intelligibility, (3) they demonstrate that the difference between hearing-impaired and normally-hearing listeners in the experiment of Best *et al.* did not simply result from hearing-impaired listeners requiring more time to prepare effectively for an upcoming talker than normally-hearing listeners, and (4) they demonstrate that atypical preparatory EEG activity accompanies poorer speech intelligibility. In combination, these findings are consistent with the explanation that atypical preparatory attention is one factor that contributes to difficulties communicating in multi-talker listening environments. Importantly, these results suggest that differences in *transduction* at the auditory periphery may not fully explain differences in speech intelligibility between normally-hearing and hearing-impaired participants.

7.3.5. Methodological contributions

In seeking significant differences, the EEG experiments in this thesis employed Spatio-temporal Cluster-based Permutation Analyses. Although this method incorporates no *a priori* assumptions, consistent results were observed across different groups of participants of different ages. This result demonstrates the reliability of the technique and its usefulness for future experiments, particularly if the electrodes or time points at which an effect may occur are not known in advance. The method corrects for multiple comparisons by grouping data points over space and time – therefore, the p-values that result from this comparison do not require further correction when the max statistic is taken at each permutation, which was the method employed in this thesis.

7.4. Future research

7.4.1. A method to eliminate differences in arousal between the Test and Control Conditions

The EEG experiments in this thesis showed similar activity on Location and Gender trials when ERPs were compared between the Test and Control Conditions, although the causes of this domain-general activity are unclear. One possibility is that differences between the Test and Control Conditions reflect differences in attention, because the visual cues had implications for attention in the Test but not in the Control Condition. However, an alternative possibility is that differential activity between the Test and Control Conditions reflected differences in arousal, because the task was more difficult in the Test Condition.

A future experiment is desirable to identify activity that relates specifically to attentional processing, which cannot be explained by differences in arousal. One possible experiment would have a similar design to the experiments in this thesis, but implement a Control Condition in which the task was more difficult for participants than the Control Condition used in the experiments of this thesis. Ideally, the accuracy of speech intelligibility would be equal (and also below ceiling level) for the Test and Control Conditions. For example, the visual cues could be degraded in the Control Condition, thus increasing the difficulty of discriminating between the cue types. In this possible experiment, differential activity between the Test and Control Conditions that was similar to the EEG activity reported in the experiments of this thesis could not be attributed to arousal, since the Test and Control Conditions would not differ in accuracy.

7.4.2. A method to eliminate the effects of differences in visual cues between conditions

One aim of this thesis was to produce a task of cued multi-talker listening suitable for children; therefore, the visual cues were designed to be intuitive and require little learning. However, one consequence was large differences in physical aspects of the visual cues presented in location and gender conditions. This aspect of the design led to a series of analyses that attempted to test the hypothesis that differences in physical aspects of the visual cues resulted in the observed differences in ERPs. These analyses involved second order comparisons and were perhaps overly-stringent for detecting significant differences.

Previous experiments measuring brain activity during preparatory attention (Hill & Miller, 2010; Lee et al., 2013) have presented visual cues with smaller physical differences between conditions than the current experiments. However, these previous experiments were not rigorous in ruling out the explanation that even small differences in the visual cues (e.g. chevron orientation) might have contributed to differences in brain activity. Since differences in the visual cues produced differences in EEG activity in the current experiment, it would be interesting to investigate whether differences in the physical structure of the cues presented by Hill and Miller and Lee *et al.* also produce differences in brain activity. A control condition could be employed with a similar design as the one used in this thesis. The results would help to distinguish between aspects of activity in their experiments that was triggered by physical and cognitive processing of the visual cues and aspects of activity that reflected preparatory attention for the location or F0 of a target talker.

Ideally, in future experiments, the stimuli used to cue attention should be counterbalanced between conditions. For example, the stimuli used to cue attention to left and right locations for one group of participants could be used to cue attention to male and female talkers (or high and low F0s) in the other group of participants. This aspect of the design would ensure that any differences between cueing conditions were due to attentional activity evoked by the visual cues, rather than physical aspects of the visual stimuli that differed between conditions.

7.4.3. A method to investigate the causes of cue-specific activity

Cue-specific activity in the experiments of this thesis could have resulted from either quantitative or qualitative differences in activity when participants attended to the location or gender of a talker. One possibility is that one of the conditions evoked a greater magnitude of activity in one region of the brain than the other condition, thus leading to differences in the amplitude of ERPs at the scalp. However, an alternative possibility is that the two conditions activated different areas of the brain.

In order to disambiguate these alternatives, Chapter 5 illustrated the results of source reconstruction of the EEG data. However, this experiment was not able to determine the sources of EEG activity with statistical robustness. Therefore, those results should be interpreted with that limitation in mind.

A future experiment could aim to test whether differences in scalp activity reflected a qualitative or quantitative difference by aiming to identify the neural generators of EEG activity. One method by which the accuracy source reconstruction could be improved is by measuring the positions of electrodes on each participant's head using electromagnetic tracking and digitisation methods (Koessler et al., 2007). This method might lead to more accurate and reliable source localisation results across participants. Another possible method would be to record EEG and fMRI activity simultaneously, thus allowing conclusions about both the timing and location of brain activity when participants attended to the location and to the gender of a talker. Gaining more accurate estimates of the locations of source activity would disambiguate whether participants utilised activity in different areas of the brain during location and gender trials or whether activity in the same areas contributed to activity in both conditions, but with higher magnitude in one condition than the other.

7.4.4. Individual differences in normally-hearing adults

Previous experiments have found large individual variation in the accuracy of speech intelligibility in the presence of other sounds (e.g. Lutfi, Kistler, Oh, Wightman, & Callahan, 2003). The experiments reported in this thesis aimed to identify EEG activity that was consistent in time and neural generator location across participants. However, it is possible that other aspects of brain activity differed in time and/or amplitude between participants, which might partially account for individual variability in speech intelligibility. For normally-hearing participants, the tasks in this thesis were too easy to observe large variations in accuracy between participants. For hearing-impaired children, who showed large differences in accuracy, the sample size was too small to provide enough statistical power for investigating individual differences.

Consistent with the hypothesis that individual differences in the timing or amplitude of EEG activity contribute to differences in behavioural performance, Choi, Wang, Bharadwaj, and Shinn-Cunningham (2014) recently reported that differences in the amplitude of the N1 potential correlated with selective attention ability. They presented three different melodies to participants at different spatial locations (left, right, and centre). Each melody contained three or four harmonic complex tones. A cue for left or right indicated the target melody for each trial and participants had to identify the contour of the target melody (ascending, descending, or zigzagging). The results showed a significant correlation: Participants who showed the largest difference in the N1 response to the left and right melodies were most accurate at identifying the melody contour. Therefore, these results demonstrate a link between the amplitude of EEG activity and performance on melody contour identification. However, it is currently unclear whether differences in EEG activity also correlate with the accuracy of speech intelligibility.

A similar experiment as reported in Chapter 5 could measure brain activity in a larger group of normally-hearing participants on a more challenging task that shows greater individual variability in speech intelligibility. There are several ways in which the task could be made more difficult for normally-hearing adults. For example, (1) the perceptual load could be increased by adding a forth distracting talker, (2) the attentional load could be increased by adding a concurrent visual task, or (3) the discriminability of the talkers could be decreased by decreasing the spatial or F0 separation between talkers. This experiment could be performed while brain activity is measured using fMRI and EEG in order to investigate whether the location or timing of brain activity (respectively) predicts the accuracy of speech intelligibility. Multi-variate regression methods could be used to identify aspects of brain activity in which the variance is associated with variability in the accuracy of speech intelligibility. The proposed experiment would lead to improved understanding of the brain activity that leads to better or worse speech intelligibility performance in normally-hearing listeners. Ultimately, this knowledge might have the potential to improve understanding of the factors that contribute to poor overall accuracy or wide variability of speech intelligibility in hearing-impaired listeners.

7.4.5. Confirming difficulties with preparatory attention in hearing-impaired children

One key question is whether atypical preparatory attention in hearing-impaired children also generalises to everyday multi-talker listening. The task presented in this thesis provided instructive visual cues to direct attention to the location or gender of an upcoming talker. In everyday life, an equivalent situation might arise when a listener sees someone to whom they wish to listen and knowledge, from vision, of their location or gender helps them to hear out the talker's voice amongst several competing voices. Therefore, the processes tested in this task are likely similar to those that would occur during everyday multi-talker listening.

One possible argument, however, is that the task was more difficult than in everyday life, which might have caused lower accuracy of speech intelligibility in hearing-impaired children. For example, words from the possible response set of colour-number combinations were spoken by all three of the talkers and the colour and number words were temporally aligned between the talkers; whereas, in everyday life, words spoken by different talkers are more likely to differ semantically and/or temporally. Taking this argument one step further, it is possible that hearing-impaired children direct preparatory attention successfully in everyday life, but did not do so in this task because they achieved low accuracy of speech intelligibility.

To test this idea, a future experiment could compare EEG activity between hearing-impaired and normally-hearing children when the accuracy of speech intelligibility was equal for the two groups. The design could be similar to that used in Experiments 3 and 4 of Chapter 6. The task could be made more difficult for normally-hearing children by degrading the acoustical stimuli, presenting white noise simultaneously with the talkers, or decreasing the angle of spatial separation in the azimuth plane between the left and right loudspeakers. In addition, the task could be made easier for hearing-impaired children by increasing the spatial separation between the left and right loudspeakers or by presenting only two talkers simultaneously. If the ERPs evoked during the Preparatory Phase of the task differ between normally-hearing and hearing-impaired children when the accuracy of speech intelligibility is equal, then this result would demonstrate differences in the brain activity used for preparatory attention, even at equivalent levels of behavioural performance. This finding would be consistent with the idea that children with moderate hearing loss have difficulties with preparatory attention when they achieve high accuracy of speech intelligibility.

7.4.6. Improving auditory attention in hearing-impaired listeners

The results reported in Chapter 6 have possible implications for the rehabilitation of children with hearing loss. The results suggest that differences between normally-hearing and hearing-impaired listeners are not limited to the periphery. Hearing-impaired listeners showed atypical preparatory attention, which was not altered when listeners used their acoustic hearing aids. This result suggests that acoustic hearing aids do not fully compensate for the difficulties faced by hearing-impaired listeners during multi-talker listening. Therefore, rehabilitative audiology should also consider ways in which attention can be improved with the aim of improving communication in noisy environments.

There have been mixed previous findings on whether auditory or cognitive training improve the accuracy of speech intelligibility in noise (e.g. Burk & Humes, 2007; Song et al., 2012). However, even in previous studies that report an improvement in performance following training, it is often not clear which aspect of training leads to an improvement in performance—first, because training often includes many diverse tasks, and second, because the improvement in performance following training is often compared with a control group who do not undertake any training, rather than a control group who undertake a different set of training that lacks one critical aspect of the experimental training. Together, these shortcomings make it difficult to attribute improvements in speech intelligibility to changes in any particular ability.

Given that this thesis shows atypical preparatory attention in hearing-impaired participants, future training studies could be aimed specifically at targeting improvements in preparatory attention. One possible training procedure could begin with a two-talker task in which the spatial and F0 separation between talkers is high. Participants would be cued to one talker on each trial and receive feedback on each trial about the correct response. The aim would be to commence training by presenting a task that hearing-impaired listeners can perform relatively accurately. At each stage, once accuracy surpasses a pre-specified criterion, the task could become progressively more difficult—for example, by decreasing the spatial or F0 separation between talkers or by presenting a greater number of competing talkers.

A control group could undergo the same duration of training on a similar task, but without preparation time. For example, they could listen to the same stimuli and perform the same task, but the target talker could be determined by a key word within the sentence, such as a particular call sign within one of the spoken sentences, rather than an instructive visual cue. The rationale is that the control group would experience identical acoustical stimuli as the experimental group, but would not experience advance cues that direct preparatory attention to one talker. All participants would be tested on the same multi-talker listening task prior to and following training, which would involve different stimuli to the training stimuli (including a different set of talkers speaking different sentences). Larger improvements for the experimental than control group (when comparing performance following training as the same task prior to training) could be attributed to improvements in preparatory attention.

In addition, following training, both groups could be tested on the task that the control group performed during training (in which preparatory attention for location or gender would not benefit speech intelligibility). If the experimental training group did not achieve improved intelligibility above the control group, this result would suggest that the benefit of training specifically improved intelligibility under circumstances in which preparatory attention had the potential to improve speech intelligibility. Therefore, the effect of training could be attributed specifically to improved preparatory attention.

7.4.7. Atypical attention in acquired hearing loss

Chapter 6 provides evidence for atypical preparatory attention in children with early-onset hearing loss. Children with early-onset hearing loss might have never learned to deploy preparatory auditory attention because they have received a distorted input from their ears for the majority of their lives. However, it is currently unclear whether equivalent degrees of hearing loss that are acquired later in life also cause difficulties preparing attention for multi-talker listening. One hypothesis is that patients who have acquired hearing loss post-lingually have experience attending to speech prior to their hearing loss and, therefore, have preserved preparatory attention. However, an alternative hypothesis is that, following a distorted input from the ear as a consequence of hearing loss, these listeners no longer have access to the acoustical information required to attend selectively to a talker based on the talker's location or gender and, therefore, lose the ability to successfully deploy preparatory attention.

In order to test the hypothesis that adults with acquired hearing loss show atypical preparatory attention, EEG activity could be measured during multi-talker listening in young adults who have acquired hearing loss. A control group would consist of normally-hearing adults who are age-matched to the hearing-impaired group. The same design could be used as Experiment 4 in Chapter 6. Comparing EEG activity between adults with normal-hearing and acquired hearing loss would reveal whether adults with acquired hearing loss show atypical preparatory attention. In addition, comparing these results with the results from hearing-impaired children reported in Chapter 6 would reveal whether early- and late-onset hearing loss have similar or different consequences for EEG activity during preparatory attention. If differences are found between individuals with early and late-onset hearing loss, then this might have different implications for the effective rehabilitation of those patients.

7.4.8. Application to other populations

The technique for identifying atypical preparatory attention described in this thesis has the potential to be applied more widely beyond listeners with peripheral hearing loss. For example, patients diagnosed with auditory processing disorder (APD) typically report listening difficulties, despite normal audiograms (Moore, Rosen, Bamiou, Campbell, & Sirimanna, 2013; Moore, 2014). A symptom of these listening difficulties is reduced speech intelligibility in noise (e.g. Bamiou, Musiek, & Luxon, 2001). A large debate surrounds whether listening difficulties result from impaired bottom-up processing of acoustical stimuli or impaired cognition (e.g. Moore et al., 2013; Moore, 2014).

The technique described in Chapter 6 has the potential to be applied to participants who have been diagnosed with APD to detect atypical preparatory attention, which is one aspect of central processing that might be impaired. Importantly, although only preparatory attention would be measured (rather than more comprehensive tests of central processing), the advantage is that atypical preparatory attention could be detected without confounding possible differences in bottom-up processing of acoustical stimuli. In combination with experiments investigating other aspects of peripheral and central auditory processing, the proposed experiment could lead to better definitions of listening difficulties that are faced by patients who have been diagnosed with APD.

7.4.9. Overall conclusions

Normally-hearing adults achieved improved speech intelligibility when they had longer durations of time to prepare their attention for the location or gender of a target talker during multi-talker listening. During preparation, EEG revealed similar spatio-temporal patterns of activity when participants knew information about a target talker’s location or gender. This result indicates a highly-overlapping, ‘domain-general’ network of brain activity during preparatory attention. For normally-hearing listeners, preparatory brain activity began early after participants were cued to location or gender and was sustained until the talkers began to speak.

In contrast, children with moderate cochlear hearing loss displayed atypical preparatory EEG activity when they were cued to the location or gender of an upcoming talker. This finding suggests that, in addition to distorted peripheral transduction for acoustical stimuli, children with hearing loss also experience atypical preparatory attention during multi-talker listening. Difficulties with preparatory attention might be one factor that contributes to poorer speech intelligibility in noisy environments. The implication of this finding is that acoustic hearing aids might not have the potential to restore normal processing of acoustical stimuli in hearing-impaired listeners. Future research should address the consequences of impaired central processing for effective rehabilitation.

Appendix A

Effect of Age in Children during Two-talker Listening

This appendix examined whether there was a correlation between a child's age and the accuracy of speech intelligibility or the amplitude of event-related potentials (ERPs) in the two-talker experiment reported in Experiment 3 of Chapter 3. This experiment recruited children aged 7–13 years. Within this age range, it is possible that the effects reported in Experiment 3 reflected patterns of results that were only displayed by the older children, rather than patterns that were consistent across all of the children who participated.

Previous experiments suggest that there is a relationship between the amplitude of ERPs and a child's age. For example, Bishop, Anderson, Reid, and Fox (2011) measured ERPs in response to pure-tone stimuli in two groups of normally-hearing children—one who participated when they were aged 7 and 9 years and the other who participated when they were aged 9 and 11 years. The results showed significant changes in the amplitude of the P1 and N1b ERP components between 7 and 9 and between 9 and 11 years. Therefore, differences in age have the potential to modulate evoked responses to acoustical stimuli. The current analyses aimed to investigate whether the accuracy and ERP results reported in Chapter 3 varied within the age group tested.

A.1. Methods

A.1.1 Participants, Apparatus, Stimuli, and Procedure

The participants, apparatus, stimuli, and procedure are reported in Experiment 3 of Chapter 3.

A.1.2 Behavioural Analyses

The analyses included Location (left/right cues) and Gender (male/female cues) trials for the Test Condition. Responses were scored as correct if both the colour and number key words were reported correctly. For each participant, the average accuracy was collapsed across Location and Gender trials.

A.1.3 Analyses of ERPs

The average amplitude of activity in each cluster from normally-hearing children (reported in Experiment 3 of Chapter 3)—averaged over the space-by-time points in the cluster—was calculated for each participant. For Type-I clusters, amplitudes were calculated for the Test Condition. For Type-II clusters, amplitudes were calculated separately for Location and Gender trials.

A.2. Results

Pearson's product moment correlations were calculated between age and average accuracy and between age and average amplitude in each of the Type-I and Type-II clusters identified in Experiment 3 of Chapter 3 (Figures 3.10 and 3.11; Clusters 25-31).

A.2.1 Accuracy of Speech Intelligibility

Figure A.1 visualises the accuracy of speech intelligibility for individual children. The correlation between age and accuracy was not significant [$r = 0.29$, $p = 0.17$], although Figure A.1 shows a slight trend towards higher accuracy with increasing age.

A.2.2 Amplitude of ERPs

Bonferroni-corrected correlations between age and the amplitude of ERPs in Type-I clusters (during the Test Condition) and in Type-II clusters (separately during Location and Gender trials) showed no significant correlations between age and average amplitude (number of correlations performed = 9; $p \geq 0.39$).

A.3. Discussion

There were no significant correlations between age and overall accuracy or between age and the amplitude of ERPs within the clusters reported in Experiment 3 of Chapter 3. The results suggest that the findings reported in Experiment 3 of Chapter 3 were consistent across the age range that was tested.

The results differ from those of Bishop *et al.* (2011), who found significant differences in the amplitude of the P1 and N1b components in response to pure-tone stimuli in children aged between 7 and 11 years. However, the current results resulted from an analysis of ERPs at different latencies and for different stimuli as those analysed by Bishop *et al.* For the current analyses, amplitudes within each cluster were averaged across the durations of the cluster, which always started later than 70 ms after the start of each phase (and, therefore, later than 70 ms after the onset of visual or acoustical stimuli) and lasted more than 300 ms. Therefore, the time over which the ERPs were analysed was, on average, later after stimulus onset than the time at which Bishop *et al.* analysed ERP amplitudes (in which the components of interest started and finished earlier than 150 ms after the onset of the pure-tone stimuli that they presented). In the current experiment, the latencies of interest were those that showed significant differences between the Test and Control Conditions or between Location and Gender trials, as reported in Experiment 3 of Chapter 3.

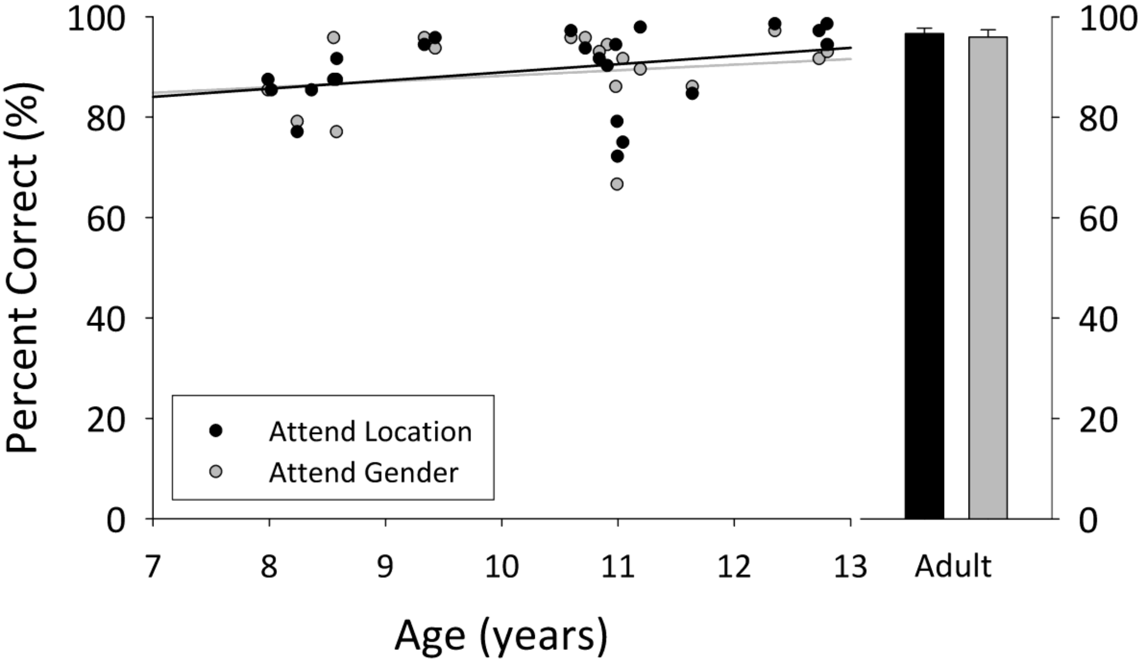


Figure A.1. Accuracy for correctly identifying the colour-number combination spoken by the target talker in the Test Condition of the EEG experiment. Accuracy for individual children who participated in Experiment 3 of Chapter 3 are each shown on the left portion of the graph. Black circles represent an individual child's accuracy on attend-location trials and grey dots on attend-gender trials. The solid black (attend-location) and grey (attend-gender) lines show the result of the best fitting least-squares linear regression equation for each condition. The bars on the right portion of the graph show average accuracy for the adults who participated in Experiment 2 of Chapter 3.

One possible caveat is that the correlation analyses may have been underpowered by the small sample size. However, correlations between age and the amplitude of ERPs all had *p*-values that were greater or equal to 0.39. Therefore, even with a larger sample of children of the same age range, it is unlikely that significant correlations would be observed for the clusters of ERPs that were analysed in this Appendix.

Overall, the results suggest that the findings reported in Experiment 3 of Chapter 3 reflected patterns of results that were consistent across all of the children who participated, rather than patterns of results that were only present in a sub-set of older or younger children.

Appendix B

Effect of Age in Children during Three-talker Listening

The aim of this appendix was to examine individual variability within the experiments reported in Chapter 6. One goal was to establish whether the average results accurately represented the results of individual listeners. Previous experiments report wide variability in performance on tasks requiring hearing-impaired listeners to attend to sounds in background noise. For example, Grose and Hall (1996) asked participants to identify a melody that was presented simultaneously with two competing melodies—one containing higher-frequency tones than the target melody and the other containing lower-frequency tones. The participant first heard the target melody in quiet and subsequently had to identify which of two intervals contained the target melody. The experiment employed an adaptive procedure in which the frequency separation between the target and competing melodies decreased after three consecutive correct responses. The procedure converged on the frequency separation that produced 79.4% correct identification performance. The results showed high variability in performance for listeners with mild-to-moderate hearing loss—some listeners required approximately 22 semitones of frequency separation, whereas others required approximately 32 semitones. These results demonstrate that large differences in performance can arise within a group of hearing-impaired listeners with similar audiograms. In the current experiments, if hearing-impaired listeners varied widely in their performance, then it is possible that a different pattern of results might be observed for each individual participant. As a result, the average data would not be representative of the majority of the participants.

Another reason for examining individual variability was that the experiments reported in Chapter 6 recruited children within a broad age range—between 7 and 16 years.

This broad age range does not necessarily reflect a homogenous group, even for normally-hearing children. For example, Bishop, Anderson, Reid, and Fox (2011) measured event-related potentials (ERPs) in response to pure-tone stimuli in two groups of normally-hearing children—one who participated when they were aged 7 and 9 years and the other who participated when they were aged 9 and 11 years. The results showed significant effects of age on the amplitude of the P1 and N1b ERP components. Therefore, differences in age have the potential to modulate evoked responses to acoustical stimuli. The current analyses aimed to investigate whether differences in age were manifest as differences in the amplitude of ERPs during three-talker listening.

This appendix addressed the results of each of the four experiments from Chapter 6, in turn, and examined the roles of age and individual variability.

B.1. Ability to benefit from preparation time

The analyses reported in this section aimed to investigate individual variability in the patterns of accuracy and latency of speech intelligibility from increasing the duration of preparation time—first, for the normally-hearing children reported in Experiment 1 of Chapter 6, and second, for the hearing-impaired children reported in Experiment 2 of Chapter 6.

B.1.1 Methods

B.1.1.1 Participants, apparatus, stimuli, and procedure

The participants, apparatus, stimuli, and procedure for normally-hearing and hearing-impaired children are reported in Experiments 1 and 2 of Chapter 6, respectively.

B.1.1.2 Analyses

Trials were separated into attend-location (average left/right cues) and attend-gender (average male/female cues) groups, separately for each of the three cue-target intervals. Responses were scored as correct if both the colour and number key words were reported correctly. In addition, reaction times (RTs), measured from the onset of acoustic stimuli, were averaged over trials in which participants correctly identified the colour-number combination. RTs beyond two standard deviations from the mean for each participant were excluded from the analysis.

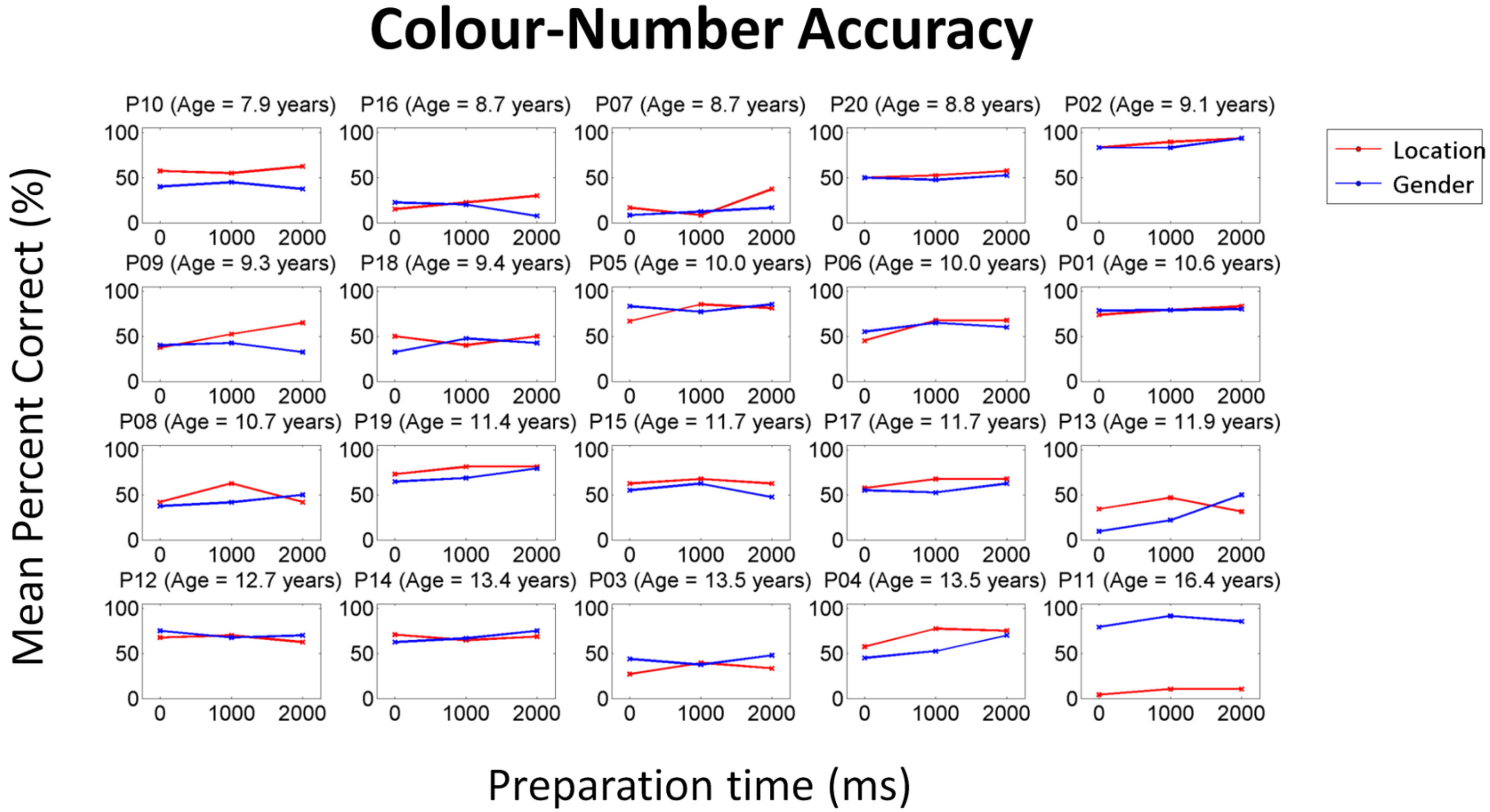


Figure B.1. Normally-hearing children: Colour-number accuracy at different durations of preparation time for individual participants. Graphs are ordered by ascending age on the top row from left to right and, subsequently, on the second, third and fourth rows.

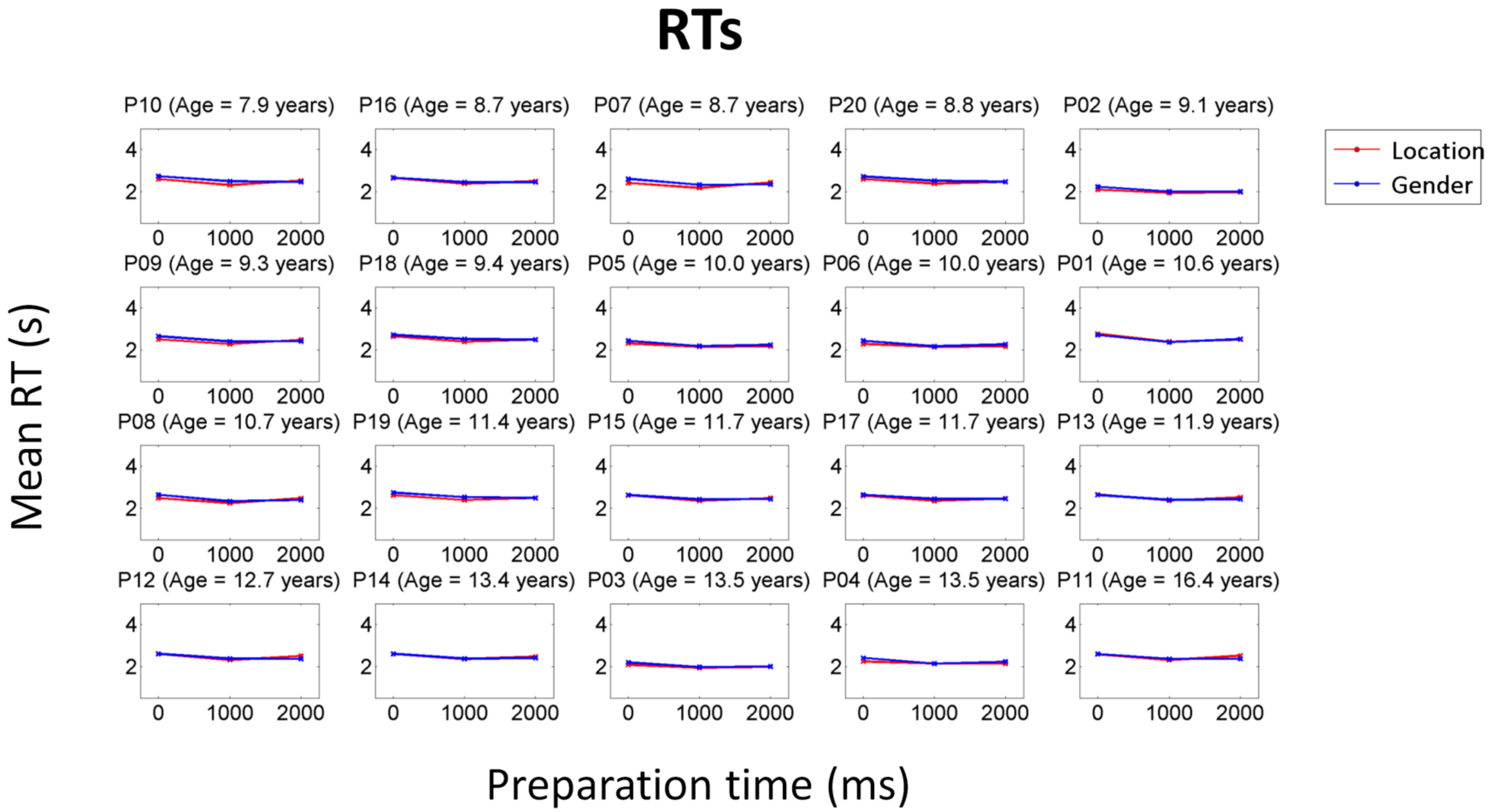


Figure B.12. Normally-hearing children: Reaction times (RTs) at different durations of preparation time for individual participants. Graphs are ordered by ascending age on the top row from left to right and, subsequently, on the second, third and fourth rows.

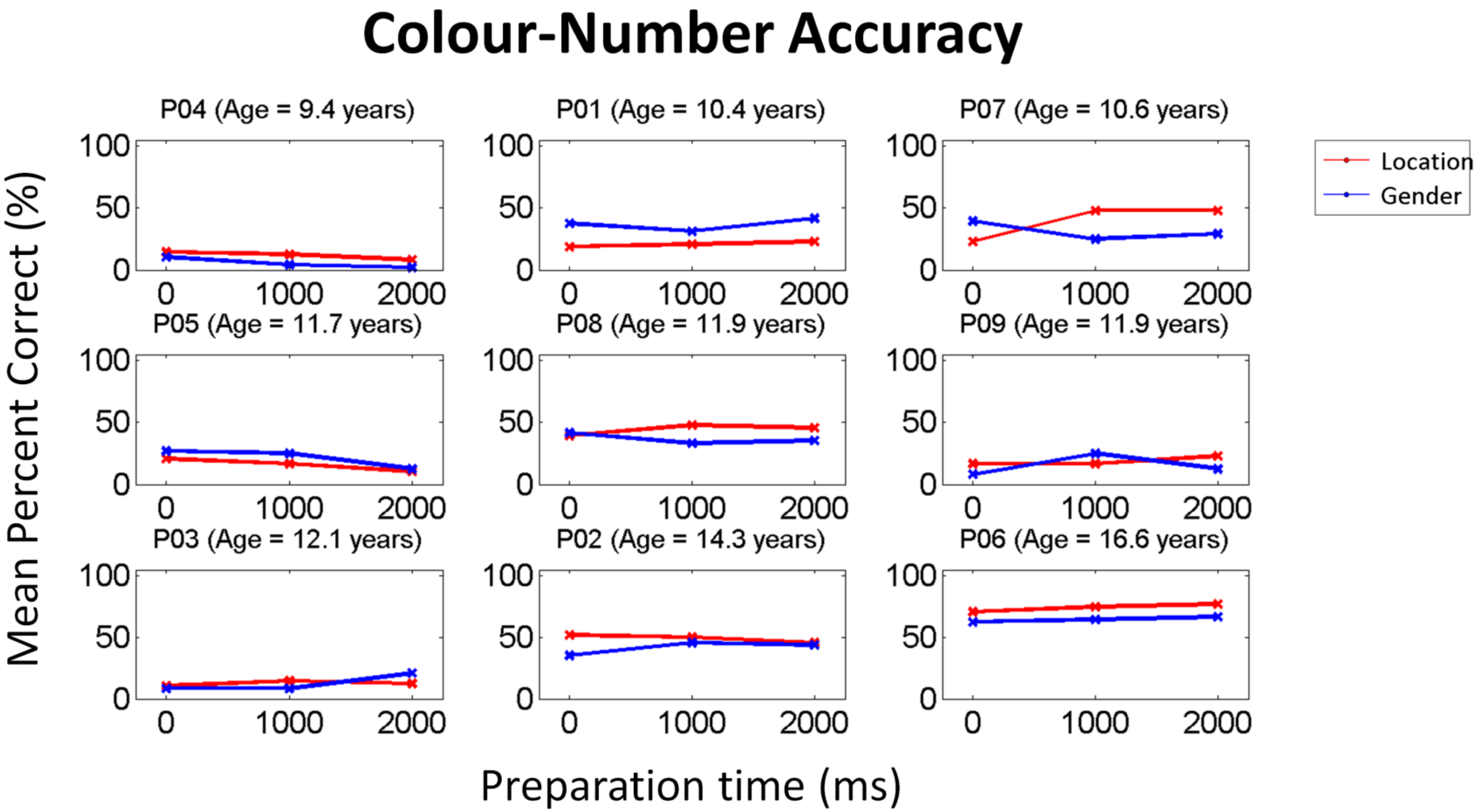


Figure B.3. Hearing-impaired children: Colour-number accuracy at different durations of preparation time for individual participants. Graphs are ordered by ascending age on the top row from left to right and, subsequently, on the middle and lower rows.

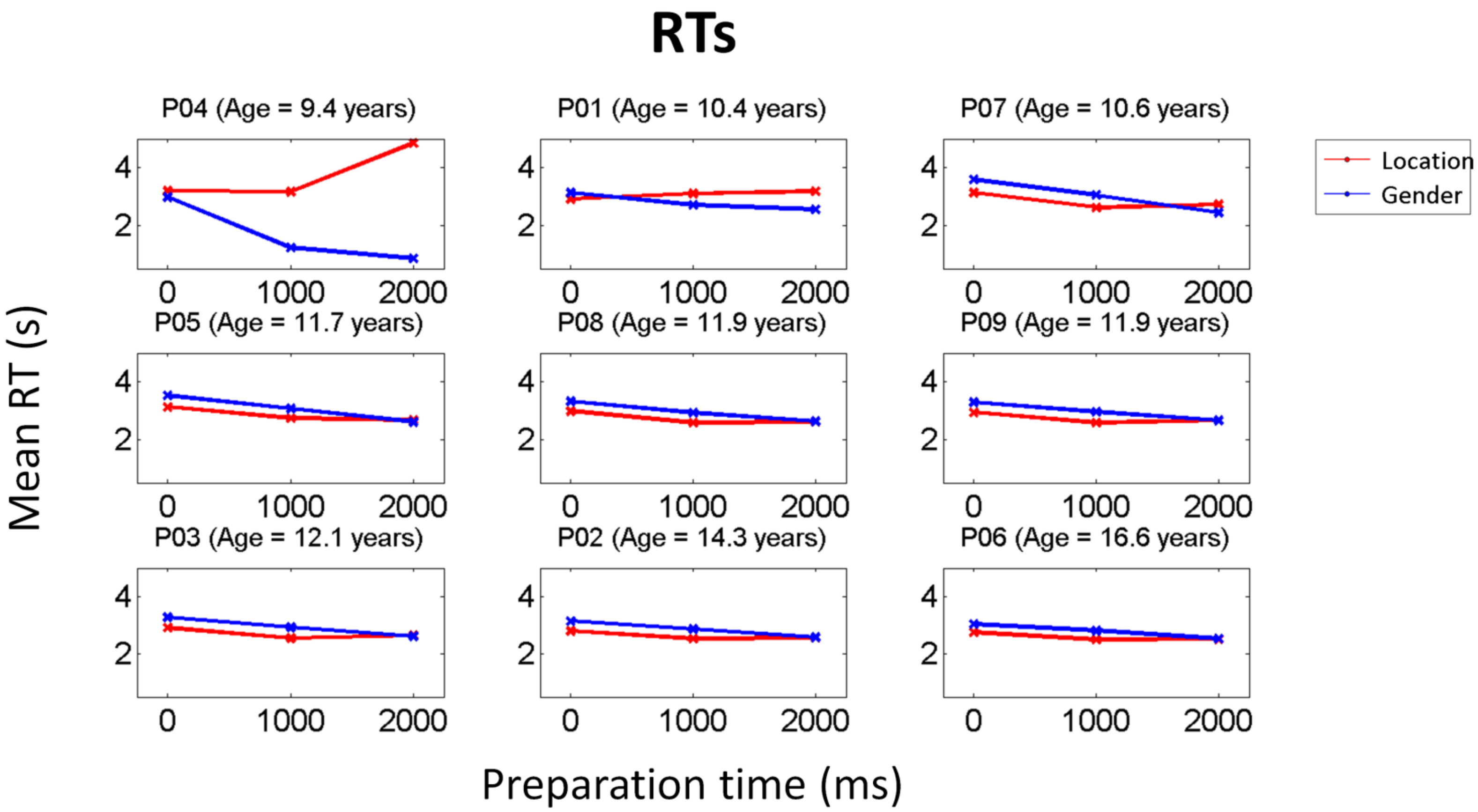


Figure B.4. Hearing-impaired children: Reaction times (RTs) at different durations of preparation time for individual participants. Graphs are ordered by ascending age on the top row from left to right and, subsequently, on the middle and lower rows.

Table B.1. Summary of Pearson's product moment correlations for normally-hearing and hearing-impaired children in Experiments 1 and 2 of Chapter 6. A tick in the second or third column indicates that the *p*-value was below the 0.05 criterion (*p*-values displayed underneath).

Factor	Significant correlation with age in normally-hearing children?	Significant correlation with age in hearing-impaired children?
Accuracy	✗ (p = 0.53)	✓ (p = 0.005)[10]
RTs	✗ (p = 0.62)	✗ (p = 0.27)

B.1.1 Results

Figures B.1 and B.2 visualise the accuracy and latency, respectively, of speech intelligibility for each of the 20 normally-hearing children who participated in Experiment 1 of Chapter 6. Figures B.3 and B.4 visualise the accuracy and latency of speech intelligibility for each of the nine hearing-impaired children who participated in Experiment 2 of Chapter 6.

Separately for normally-hearing (Experiment 1 of Chapter 6) and hearing-impaired (Experiment 2 of Chapter 6) children, accuracy and RTs were collapsed over attend-location and attend-gender trials and over the three cue-target interval conditions. Pearson's product moment correlations were performed between age and average accuracy and between age and average RTs. Table B.1 lists the results of the analyses. For normally-hearing children, there was no significant correlation between age and accuracy [$r = 0.15$, $p = 0.53$] or RTs [$r = 0.12$, $p = 0.62$]. For hearing-impaired children, there was a significant improvement in accuracy with increasing age [$r = 0.80$, $p = 0.010$], but no significant correlation between age and RTs [$r = 0.41$, $p = 0.27$]. A partial correlation showed that the relationship between age and accuracy was still significant when 5-frequency average hearing levels were taken into account [$pr = 0.87$, $p = 0.005$].

B.1.2 Discussion

For normally-hearing children, there was no significant correlation between age and average accuracy or RTs (Table B.1). In addition, participants generally showed similar patterns of accuracy scores with increasing durations of preparation time compared to the group average reported in Experiment 1 of Chapter 6—most showed higher accuracy (averaged across location and gender conditions) at the 2000 ms interval than the 0 ms

---

[10] The reported p-value results from a partial correlation between age and average accuracy, taking into account 5-frequency average hearing level.

interval. Although, the pattern of results between 0 and 1000 ms and between 1000 and 2000 ms was less consistent across participants. This finding might underlie the results reported in Experiment 1 of Chapter 6—a significant increase in accuracy between the 0-ms and 2000-ms intervals, but no significant progressive improvement in accuracy between 0-ms and 1000-ms or between 1000-ms and 2000-ms. In contrast to accuracy, RTs in normally-hearing children showed higher consistency across participants. The average latency for responding was similar across participants and the pattern of RTs with increasing durations of preparation time was also similar (Figure B.2).

For hearing-impaired children, there was a significant correlation between age and overall accuracy that could not be explained by differences in hearing level, although there was no significant correlation between age and RTs. Although age affected overall accuracy, there were similar patterns of accuracy and RTs for individual participants with increasing durations of preparation time—the slopes of the lines (with increasing durations of preparation time) were relatively flat across all participants (Figures B.3–4). Overall, most of the hearing-impaired children showed little or no improvement in accuracy with increasing preparation time and a small number of participants even showed a decrease in accuracy with increasing preparation time. These results are consistent with the results reported for the group of hearing-impaired children in Experiment 2 of Chapter 6, which demonstrated no significant increase in accuracy as the duration of preparation time increased. Therefore, even though age modulated the average accuracy of hearing-impaired children, it did not affect the ability to benefit from increasing preparation time.

For the RT data, individual patterns were similar across hearing-impaired participants (Figure B.4). Apart from the youngest child that participated, who displayed an atypical pattern of RTs compared to the other hearing-impaired children, the hearing-impaired children showed responses of similar latencies. In addition, the pattern of RTs with increasing durations of preparation time was consistent—most showed progressively shorter RTs as the duration of preparation time increased. This result is consistent with the results reported in Experiment 2 of Chapter 6, which showed a significant progressive shortening of latencies with increasing durations of preparation time.

B.2. Accuracy and amplitude of ERPs during EEG experiments

The analyses reported in this section aimed to investigate individual variability and effects of age on the data reported in Experiments 3 and 4 of Chapter 6. Analyses were conducted on the average accuracy of speech intelligibility and the amplitude of ERPs.

B.2.1 Methods

B.2.1.1 Participants, apparatus, stimuli, and procedure

The participants, apparatus, stimuli, and procedure for normally-hearing and hearing-impaired children are reported in Experiments 3 and 4 of Chapter 6, respectively.

B.2.1.2 Behavioural analyses

The analyses included Location (left/right cues) and Gender (male/female cues) trials for the Test Condition. Responses were scored as correct if both the colour and number key words were reported correctly. For each participant, average accuracy was collapsed across Location and Gender trials.

B.2.1.3 Analyses of ERPs

The average amplitude of activity in each cluster from normally-hearing children (reported in Experiment 3 of Chapter 6)—averaged over the space-by-time points in the cluster—was calculated for each participant. For Type-I clusters, amplitudes were calculated for the Test Condition. For Type-II clusters, amplitudes were calculated separately for Location and Gender trials in the Test Condition.

B.2.2 Results

Pearson's product moment correlations were performed between age and average accuracy and between age and average amplitude in each of the Type-I and Type-II clusters identified in Experiment 3 of Chapter 6. Table B.2 lists the results of the analyses.

Figure B.5 visualises the accuracy of speech intelligibility for individual normally-hearing children. There was a significant improvement in accuracy with increasing age, $r = 0.42$, $p = 0.039$. Bonferroni-corrected correlations between age and the amplitude of ERPs in Type-I clusters (during the Test Condition) and in Type-II clusters (separately during Location and Gender trials) showed no significant correlations between age and amplitude (number of correlations performed [N] = 15; $p > 0.99$).

Figure B.6 visualises the accuracy of speech intelligibility for individual hearing-impaired children. There was a significant improvement in accuracy with increasing age, $r = 0.53$, $p = 0.023$. However, the correlation between age and accuracy was *not* significant when

Table B.2. Summary of Pearson's product moment correlations for normally-hearing and hearing-impaired children in Experiments 3 and 4 of Chapter 6. A tick in the second or third column indicates that the *p*-value was below the 0.05 criterion (*p*-values displayed underneath).

Factor	Significant correlation with age in normally-hearing children?	Significant correlation with age in hearing-impaired children?
Accuracy	✓ (p = 0.010)	✗ (p = 0.08)
ERP amplitude		
Cluster 1 (Test Condition)	✗	✗
Cluster 2 (Test Condition)	✗	✗
Cluster 3 (Test Condition)	✗	✗
Cluster 4 (Test Condition)	✗	✗
Cluster 5 (Test Condition)	✗	✗
Cluster 6 (Test Condition)	✗	✗
Cluster 7 (Test Condition)	✗	✗
Cluster 8 (Test Condition)	✗	✗
Cluster 9 (Test Condition)	✗	✗
Cluster 10 (Test Condition)	✗	✗
Cluster 11 (Test Condition)	✗	✗
Cluster 12 (Location trials)	✗	✗
Cluster 12 (Gender trials)	✗	✗
Cluster 13 (Location trials)	✗	✗
Cluster 13 (Gender trials)	✗	✗

5-frequency average hearing levels were taken into account [pr = 0.53, p = 0.08]. Bonferroni-corrected correlations between age and the amplitude of ERPs in Type-I and Type-II clusters showed no significant correlations between age and amplitude (N = 15; p > 0.99).

B.2.3 Discussion

In normally-hearing and hearing-impaired children, there was a significant improvement in overall accuracy with increasing age. However, after accounting for variation in average hearing level, the correlation between age and the accuracy of speech intelligibility was not significant in hearing-impaired children. There was also no significant correlation between age and the amplitude of ERPs for the clusters in which significant differences were observed in normally-hearing children (reported in Experiment 3 of Chapter 6)—either for normally-hearing children or for hearing-impaired children.

Overall, these findings are consistent with the idea that the accuracy of speech intelligibility improves with increasing age in normally-hearing children, but the brain activity that underlies the patterns of EEG activity reported in Experiments 3 and 4 of Chapter 6 is not affected by age in either normally-hearing or hearing-impaired children.

B.3. General discussion

Overall, there was some evidence that increases in age produce improvements in the overall accuracy of speech intelligibility in normally-hearing and hearing-impaired children

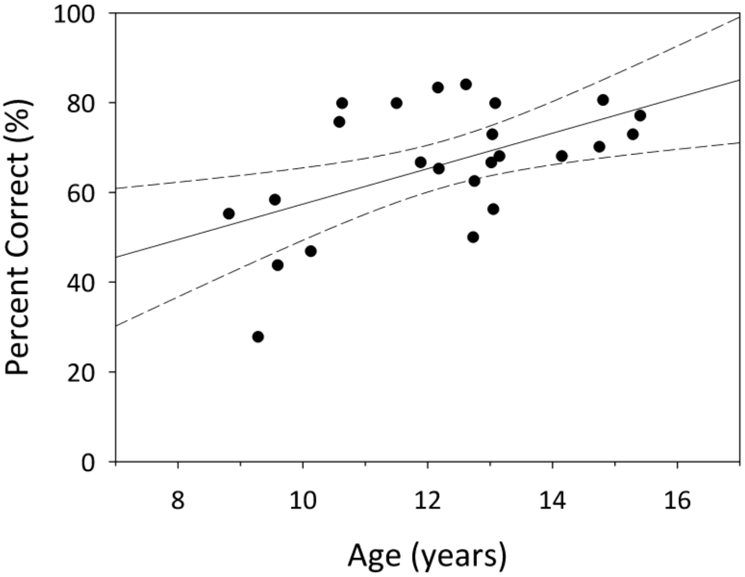


Figure B.5. Normally-hearing children: Accuracy of correctly identifying the colour-number combination spoken by the target talker in the Test Condition of the EEG experiment. Black circles each represent an individual participant. The solid black line shows the result of the best fitting least-squares linear regression equation. The dashed lines indicate upper and lower 95% confidence intervals of the regression estimates.

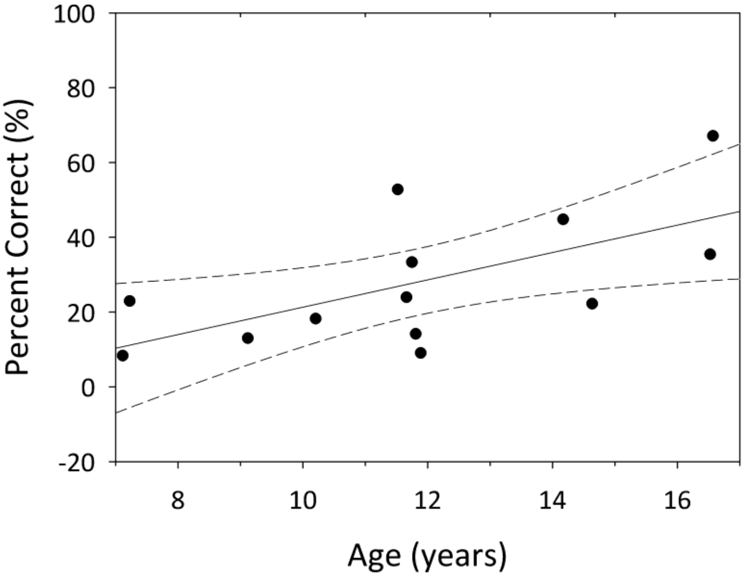


Figure B.6. Hearing-impaired children: Accuracy of correctly identifying the colour-number combination spoken by the target talker in the Test Condition of the EEG experiment. Black circles each represent an individual participant. The solid black line shows the result of the best fitting least-squares linear regression equation. The dashed lines indicate upper and lower 95% confidence intervals of the regression estimates.

(Tables B.1–2). However, age was not associated with the overall latency of responses or the benefit of longer durations of preparation time for the accuracy and latency of speech intelligibility. In addition, the EEG activity that occurred in Type-I and Type-II clusters was not significantly correlated with age. The findings that age does not affect the benefit to speech intelligibility from increasing the duration of preparation time or the amplitude of preparatory EEG activity are compatible with each other. Overall, the results suggest that the findings reported in Chapter 6 were not restricted to particular ages of participants within the broad age groups that were tested.

However, one possible limitation is that the correlations were underpowered by small samples of participants. In order to thoroughly rule out the explanation that age affected the amplitude of ERPs, a replication would be desirable with greater numbers of participants. Nevertheless, the finding of significant correlations between age and overall accuracy suggests that these analyses had sufficient power to detect large and consistent effects of age.

B.3.1 Conclusions

Overall, during the three-talker listening task, the patterns of accuracy and RTs across different cue-target intervals were consistent across children of different ages, as were the amplitudes of ERPs. These findings applied to data collected from both normally-hearing and hearing-impaired children. Therefore, the ages of participants within the age range tested were unlikely to have significantly influenced the conclusions that are reported in Chapter 6.

Appendix C

Effect of Aiding on ERPs

Experiment 4 reported in Chapter 6 revealed similarities and differences in the event-related potentials (ERPs) recorded during three-talker listening between children with normal hearing and children with moderate hearing loss. The data for the hearing-impaired children were obtained when they performed the task without their hearing aids. This appendix reports the results from an additional session of the experiment that investigated whether similar patterns of ERPs were observed when hearing-impaired participants performed the task with and without their acoustic hearing aids. This comparison aimed to rule out two possible alternative explanations for differences in ERPs between normally-hearing and hearing-impaired children: (1) inaudibility of the acoustical stimuli in hearing-impaired children, and (2) unfamiliar listening conditions in hearing-impaired children.

In more detail, ruling out these alternative explanations was important for two main reasons. First, although the acoustical stimuli were presented at a higher level for hearing-impaired than normally-hearing children (which aimed to partially compensate for differences in audiometric pure-tone thresholds), it is possible that aspects of the acoustical stimuli remained inaudible for some of the hearing-impaired children (in particular, those with poorer audiometric thresholds). Second, all of the children listened with their hearing aids in everyday life more often than they listened without hearing aids. Results showing similar patterns of amplitudes when participants performed the task with and without their hearing aids would rule out the explanation that inaudibility or unfamiliar listening conditions explained differences between normally-hearing and hearing-impaired children that are reported in Experiment 4 of Chapter 6.

This appendix reports within-subjects comparisons from the subset of nine hearing-impaired children who performed both sessions of Experiment 4—one aided and one unaided. Acoustic hearing aids have been shown to provide a small benefit on self-report measures of speech intelligibility in noisy environments (Gatehouse & Akeroyd, 2006; Noble & Gatehouse, 2006; Noble, 2006). Therefore, it was expected that hearing aids might slightly improve the accuracy of speech intelligibility or alter the ERPs evoked by the acoustical stimuli (i.e. during the Selective Phase). However, hearing aids were not expected to alter ERPs recorded during the Preparatory Phase (before the acoustical stimuli began). Of particular interest were amplitudes of the clusters that showed atypical ERPs in hearing-impaired children without their hearing aids (reported in Experiment 4 of Chapter 6)—since any conclusions drawn about differences in auditory attention between normally-hearing and hearing-impaired children rely on the assumption that inaudibility or unfamiliar listening conditions cannot explain the results.

C.1. Methods

C.1.1 Participants

Participants were 9 children (2 male), aged 7–16 years (mean [M] = 11.6, standard deviation [SD] = 2.7) who completed two sessions of the experiment: the first without their hearing aids and the second using their own bilateral behind-the-ear acoustic hearing aids. Out of the nine children, one was left-handed and had an additional visual impairment in her left eye. The results from the unaided session are reported in Experiment 4 of Chapter 6[11]. The aided session took place between 2 and 9 months after each child participated in the unaided session.

C.1.2 Apparatus, stimuli, and procedure

The apparatus, stimuli, and procedure are reported in Experiment 4 of Chapter 6.

C.1.3 EEG recording and processing

EEG recording and processing were the same as Experiment 3 of Chapter 6, except that M1 and M2 were not recorded in the aided session because the placement of the hearing aids obscured these scalp locations.

C.1.4 Analyses

C.1.4.1 Behavioural analyses

Trials were separated into Location (average left/right cues) and Gender (average male/female cues) groups, separately for the Test and Control Conditions. Responses were scored as correct if both the colour and number key words were reported correctly in the Test Condition and if the visual cue was reported correctly in the Control Condition. Percent-correct accuracy in hearing-impaired listeners was compared between aided and unaided

---

[11] Although, the results reported in Experiment 4 of Chapter 6 also include four additional participants who did not participate in the second session.

sessions. Accuracy in each session was also compared to the accuracy of normally-hearing listeners who participated in Experiment 3 of Chapter 6.

C.1.4.2 Analyses of ERPs

The first analysis compared the overall ERP waveforms measured from hearing-impaired children between sessions in which they performed the task with and without their hearing aids. First, trials were separated into those that occurred during the Test and Control Conditions. Next, amplitudes were averaged across a broad group of 28 posterior electrodes and a separate group of 34 anterior electrodes.

The second analysis compared the average amplitude of activity in each cluster from normally-hearing children in Experiment 3—averaged over the space-by-time points in the cluster—in hearing-impaired children between sessions in which they performed the task with and without their hearing aids.

C.2. Results

C.2.1 Behavioural results

A 2 x 2 within-subjects ANOVA was conducted with the factors aiding (aided/unaided) and cue type (Location/Gender), separately for the Test and Control Conditions (Figure C.1). In the Test Condition, participants achieved significantly higher accuracy when they performed the task aided (M = 39.5%, SD = 20.8) than unaided (M = 25.8%, SD = 19.6), $F(1, 8) = 27.3$, $p = 0.001$, $\eta_p^2 = 0.78$. However, there was no significant main

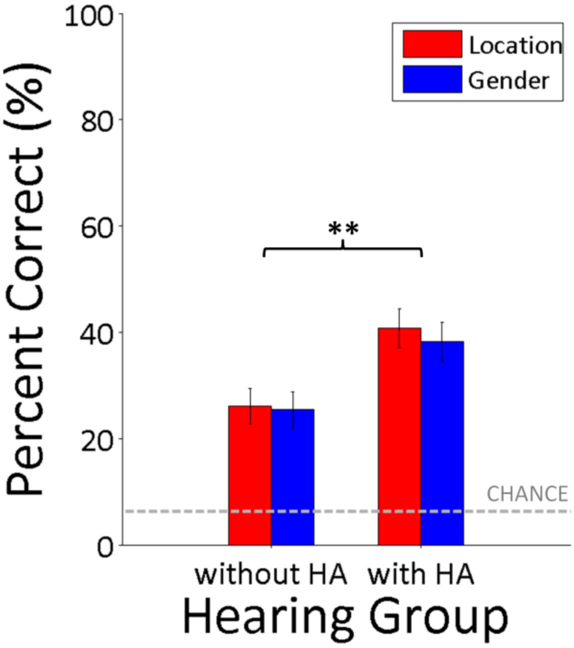


Figure C.1. Mean percentage of trials in which participants correctly identified the colour-number combination spoken by the target talker, plotted as separate bars for hearing-impaired children in Experiment 4 who completed the task with and without their hearing aids (HA).

effect of cue type [$F(1, 8) = 0.91$, $p = 0.37$] and no significant interaction between hearing aids and cue type [$F(1, 8) = 0.49$, $p = 0.50$]. In the Control Condition, there was no significant difference in accuracy between aided (M = 93.2%, SD = 11.0) and unaided (M = 93.8%, SD = 9.5) sessions, $F(1, 8) = 0.53$, $p = 0.49$. There was no significant main effect of cue type [$F(1, 8) = 0.21$, $p = 0.66$] and no significant interaction [$F(1, 8) = 4.40$, $p = 0.07$].

To compare behavioural accuracy with the normally-hearing children tested in Experiment 3, a 3 x 2 mixed ANOVA was conducted with the factors hearing group (3 levels: normally-hearing, hearing-impaired aided, and hearing-impaired unaided) and cue type (Location/Gender). In the Test Condition, there was a significant main effect of hearing group [$F(2, 43) = 24.56$, $p < 0.001$, $\eta_p^2 = 0.53$]. Contrasts showed that accuracy was significantly higher for normally-hearing children than for hearing-impaired children, both with ($p = 0.001$) and without ($p < 0.001$) their hearing aids. There was no significant effect of cue type [$F(1, 43) = 3.26$, $p = 0.08$] and no significant interaction between hearing group and cue type [$F(2, 43) = 1.11$, $p = 0.34$]. In the Control Condition, there were no significant differences in accuracy between hearing groups [$F(2, 43) = 2.22$, $p = 0.12$], no significant effect of cue type [$F(1, 43) = 1.08$, $p = 0.31$] and no significant interaction between hearing group and cue type [$F(2, 43) = 0.74$ $p = 0.48$].

C.2.2 Event-related potentials

C.2.2.1 Overall ERP waveforms

Figure C.2 displays the average ERPs (averaged across Location and Gender trials and across posterior and anterior channels) that occurred in the Test and Control Conditions for hearing-impaired children in the aided and unaided sessions. For each waveform, amplitudes at each time point were compared between aided and unaided sessions in a paired-samples *t*-test. Since the sample size was small, a $p < 0.01$ criterion was applied to the uncorrected *p*-values to estimate where differences might occur. There were no significant differences between normally-hearing and hearing-impaired children at posterior or anterior channels during the Control Condition ($p > 0.01$, uncorrected). During the Test Condition, the only significant difference occurred during the Selective Phase, between 801 and 811 ms. This significant difference occurred at 11 time points. The ERPs for children performing the task with their hearing aids generally did not trend towards the ERPs for normally-hearing children, apart from the first major peak during the Selective Phase (Figure C.2).

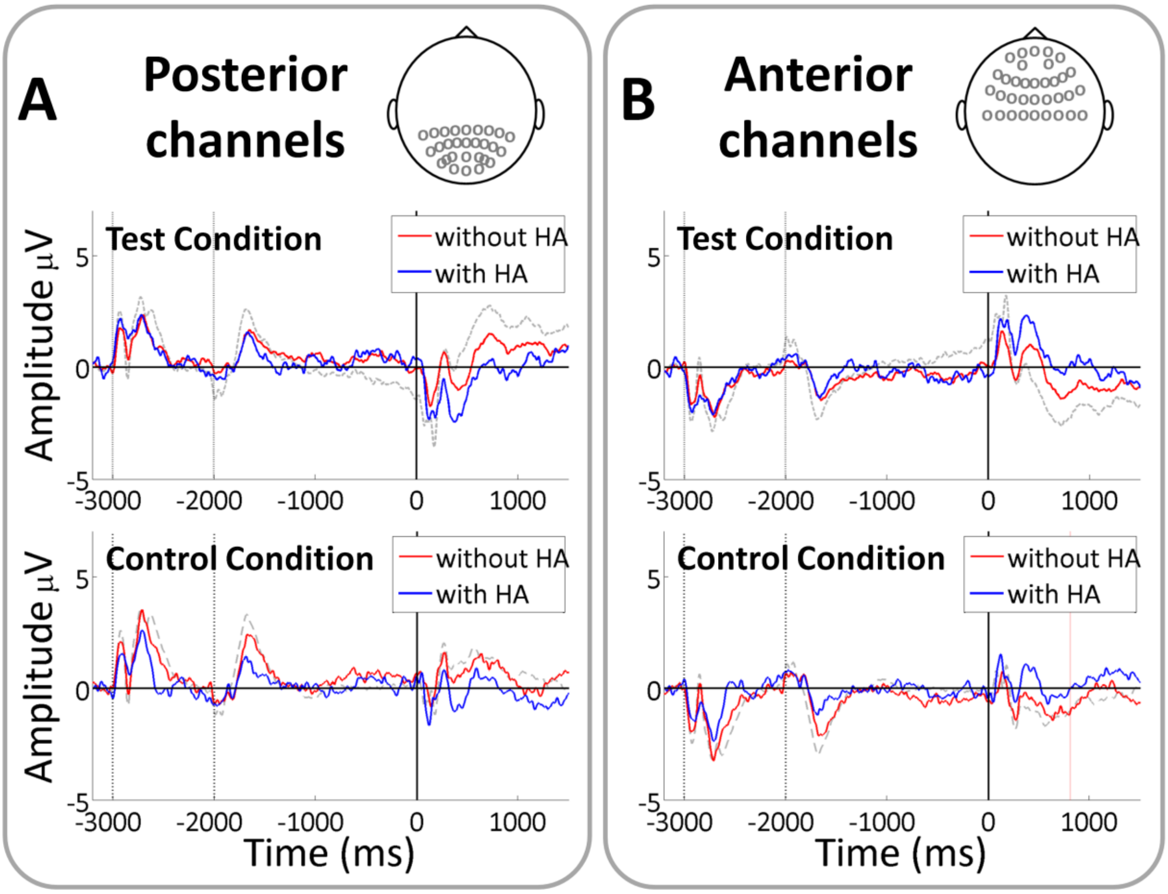


Figure C.2. Overall ERP waveforms for hearing-impaired children performing the task with and without their hearing aids (HA). The waveforms have been averaged across cue types (Location/Gender), separately for the Test and Control Conditions. Panel (**A**) displays amplitudes averaged across a group of posterior channels. Panel (**B**) displays amplitudes averaged across a group of anterior channels. The electrodes included in each average are displayed on the scalp maps at the top of each panel. For each graph, the red shaded boxes indicate time points in which an independent-samples *t*-test revealed an uncorrected *p*-value that reached the 0.01 criterion, and the grey dashed line shows the average ERP waveform for the normally-hearing children who participated in Experiment 3 of Chapter 6.

C.2.2.2 Comparisons for clusters identified in normally-hearing children

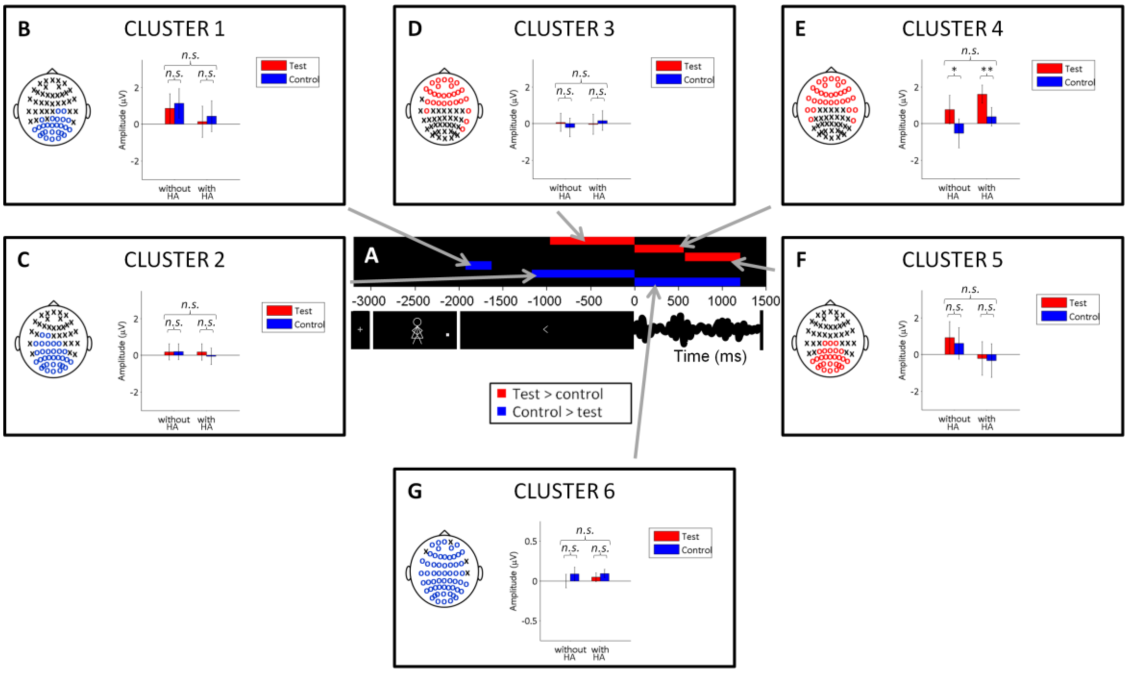
This analysis considered the amplitudes of ERPs recorded from hearing-impaired children in the significant clusters identified in the normally-hearing children who participated in Experiment 3 of Chapter 6. A paired-sampled *t*-test compared the average amplitude between the Test and Control Conditions in the aided session. If aiding did not affect the ERPs recorded from hearing-impaired children, then significant differences should only occur for the clusters in which significant differences were found when the children performed the task unaided (reported in Experiment 4 of Chapter 6). The resulting *p*-values for the Type-I clusters are tabulated in the fourth column of Table C.1. The only significant

Table C.1. Summary of within-subjects comparisons (n = 9) for hearing-impaired children between aided and unaided sessions. Amplitudes—averaged over the electrodes and time points that contribute to each cluster—were analysed for the Type-I clusters identified from normally-hearing children in Experiment 3. A tick in the column headed 'Significant difference between Test/Control Conditions?' indicates that a paired-samples t-test revealed a significant difference in amplitude between the Test and Control Conditions in the hearing-impaired children performing the task *with* their hearing aids (p-values displayed underneath). The two columns on the far right show the results of a 2 x 2 within-subjects ANOVA with the factors aiding (aided/unaided) and condition (Test/Control). A tick in the column headed 'Main effect of aiding?' indicates a significant main effect, and a tick in the column headed 'Significant interaction between hearing aid groups?' indicates that the ANOVA revealed a significant two-way interaction.

Cue Type	Phase	Cluster Number	Significant difference between Test/Control Conditions?	Main effect of aiding?	Significant interaction between aiding groups?
Location	Preparatory	1	✗ $p = 0.58$	✗ $p = 0.27$	✗ $p = 0.98$
		2	✗ $p = 0.40$	✗ $p = 0.57$	✗ $p = 0.41$
		3	✗ $p = 0.57$	✗ $p = 0.63$	✗ $p = 0.13$
	Selective	4	✓ $p = 0.004$	✓ $p = 0.030$	✗ $p = 0.88$
		5	✗ $p = 0.84$	✓ $p = 0.049$	✗ $p = 0.64$
		6	✗ $p = 0.19$	✗ $p = 0.47$	✗ $p = 0.48$
Gender	Preparatory	7	✗ $p = 0.50$	✗ $p = 0.56$	✗ $p = 0.34$
	Selective	8	✗ $p = 0.35$	✗ $p = 0.40$	✗ $p = 0.65$
		9	✗ $p = 0.10$	✗ $p = 0.06$	✗ $p = 0.88$
		10	✗ $p = 0.87$	✓ $p = 0.011$	✗ $p = 0.88$
		11	✗ $p = 0.86$	✓ $p = 0.009$	✗ $p = 0.95$

difference was for Cluster 4 ($p = 0.040$), which occurred early during the Selective Phase of Location trials. The difference between the Test and Control Conditions in Cluster 4 was also significant for hearing-impaired children *without their hearing aids* (reported in Table 6.4, Chapter 6).

Location trials



Gender trials

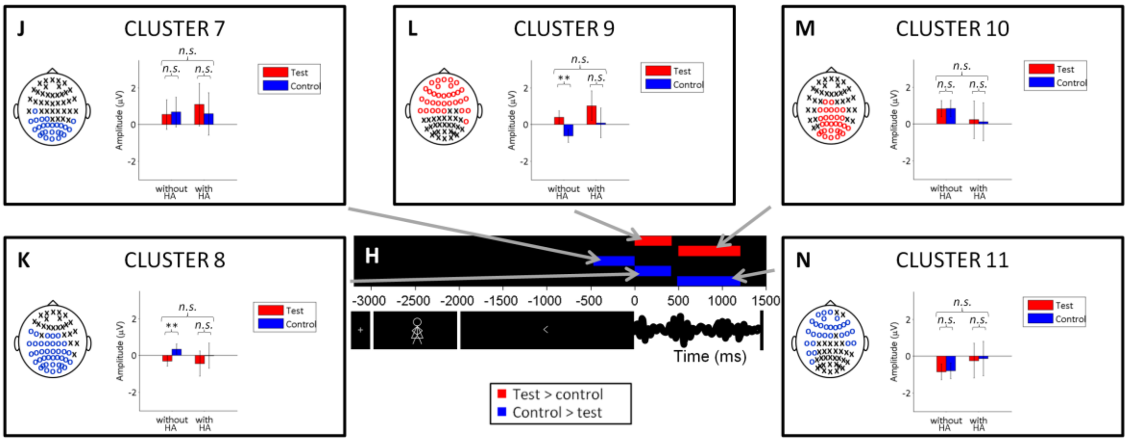


Figure C.3. Comparison of amplitudes in each cluster identified in the Type-I Spatio-temporal Cluster-based Permutation Analyses of Experiment 3 for Location (**A** to **G**) and Gender trials (**J** to **N**) in hearing-impaired children between aided and unaided sessions. (**A** and **H**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. Further information about each cluster is displayed in (**B**)-(**G**) and (**J**)-(**N**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the bar graph shows the average amplitude of ERPs—averaged over the electrodes and time points that contributed to the cluster—for hearing-impaired children performing the task with and without their hearing aids (HA; correct-and-incorrect trials). Error bars show within-subjects 95% confidence intervals. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing the Test and Control Conditions within each hearing group (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). The larger brackets at the top of each graph indicate the significance level of the interaction between hearing group and the Test/Control Conditions.

A 2 x 2 within-subjects ANOVA was conducted with the factors aiding (aided/unaided) and condition (Test/Control). There were no significant main effects of aiding for any of the Type-I clusters that occurred during the Preparatory Phase (fifth column of Table C.1). However, there was a significant main effect of aiding in two of the three clusters that occurred during the Selective Phase of Location trials (Cluster 4: *p* = 0.030; Cluster 5: *p* = 0.049) and in two of the four clusters during the Selective Phase of Gender trials (Cluster 10: *p* = 0.011; Cluster 11: *p* = 0.009). Importantly, none of the clusters showed a significant interaction between aiding and condition (sixth column of Table C.1). Figure C.3 illustrates these results.

The *p*-values for the Type-II clusters are tabulated in Table C.2. There were no significant differences between Location and Gender trials in the hearing-impaired children in the aided session, which is the same finding as previously reported in the unaided session (Table 6.5, Chapter 6). A 2 x 2 within-subjects ANOVA with the factors aiding (aided/unaided) and cue type (Location/Gender) showed no significant main effects of hearing aids and no significant interactions. Figure C.4 illustrates these results.

Table C.2. Summary of within-subjects comparisons (n = 9) for hearing-impaired children between sessions in which they performed the task with and without their hearing aids. Amplitudes—averaged over the electrodes and time points that contribute to each cluster—were analysed for the Type-II clusters identified from normally-hearing children in Experiment 3. A tick in the column headed ‘Significant difference between Location/Gender trials?’ indicates that a paired-samples *t*-test revealed a significant difference in amplitude between Location and Gender trials in the hearing-impaired children performing the task *with* their hearing aids (*p*-values displayed underneath). The two columns on the far right show the results of a 2 x 2 within-subjects ANOVA with the factors aiding (aided/unaided) and cue type (Location/Gender). A tick in the column headed ‘Main effect of aiding?’ indicates a significant main effect, and a tick in the column headed ‘Significant interaction between aiding groups?’ indicates that the ANOVA revealed a significant two-way interaction.

Cue Type	**Phase**	**Cluster Number**	**Significant difference between Location/Gender trials?**	**Main effect of aiding?**	**Significant interaction between aiding groups?**
Test	Preparatory	**12**	✗ *p* = 0.93	✗ *p* = 0.83	✗ *p* = 0.62
	Selective	**13**	✗ *p* = 0.52	✗ *p* = 0.38	✗ *p* = 0.56

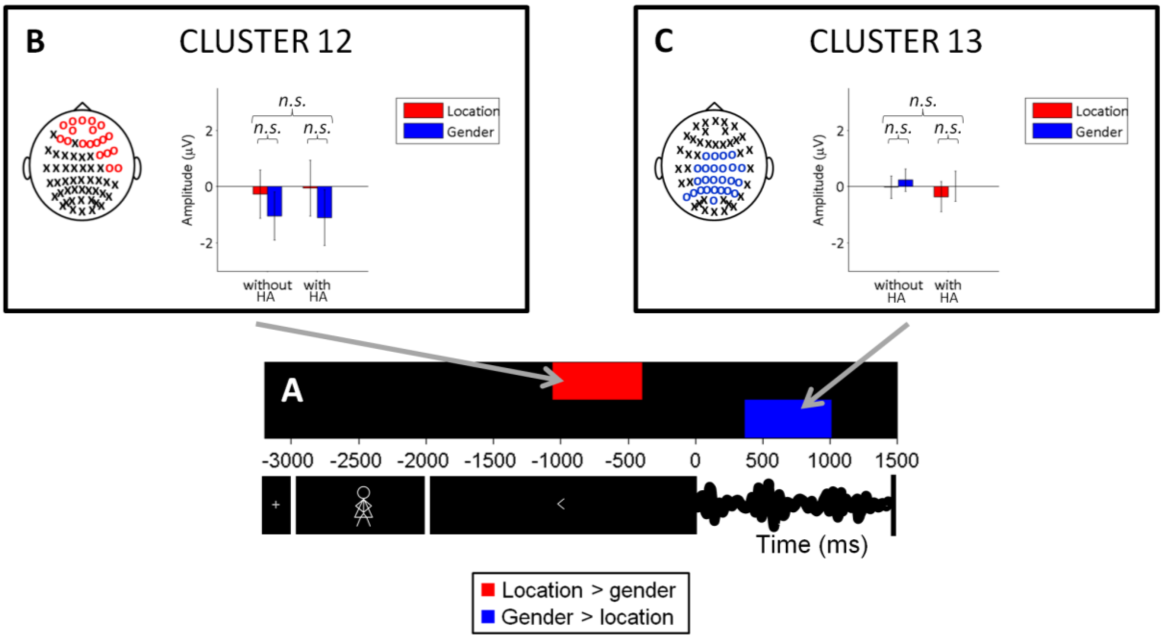


Figure C.4. Comparison of amplitudes in each cluster identified in the Type-II Spatio-temporal Cluster-based Permutation Analysis of Experiment 3, which contrasted Location and Gender trials in the Test Condition in hearing-impaired children between aided and unaided sessions. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. Further information about each cluster is displayed in (**B**)-(**C**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the bar graph shows the average amplitude of ERPs—averaged over the electrodes and time points that contributed to the cluster—for hearing-impaired children performing the task with and without their hearing aids (HA; correct-and-incorrect trials). Error bars show within-subjects 95% confidence intervals. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing the Test and Control Conditions within each hearing group (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). The larger brackets at the top of each graph indicate the significance level of the interaction between hearing group and cue type (location/gender).

C.3. Discussion

Hearing-impaired children achieved higher accuracy of speech intelligibility when they performed the task with their hearing aids than without their hearing aids. However, they still showed significantly poorer intelligibility when wearing their hearing aids than the normally-hearing children who participated in Experiment 3 of Chapter 6.

Hearing-impaired children showed similar amplitudes of ERPs when they performed the task with and without their hearing aids. First, when the overall amplitudes of ERPs were plotted for the Test and Control Conditions, only one 10-ms window showed significant differences between aided and unaided sessions (Figure C.2). Second, the clusters identified in normally-hearing children (which are reported in Experiment 3 of Chapter 6) showed similar patterns of amplitudes between aided and unaided sessions. None of the interactions between aiding and *either* condition (Type-I clusters) *or* cue type (Type-II clusters) were significant within any of the clusters (Tables C.1 and C.2). This result shows that differences in amplitudes between the Test and Control Conditions and differences in amplitudes between Location and Gender trials were not significantly affected by whether or not participants used hearing aids. Therefore, the results provide evidence that the differences between normally-hearing and hearing-impaired children reported in Experiment 4 of Chapter 6 cannot be explained by atypical listening conditions in hearing-impaired children or inaudibility of the acoustical stimuli.

The finding of similarities in ERPs when hearing-impaired children listened with and without hearing aids is consistent with previous research suggesting that hearing-impaired participants gain only a small benefit from using their hearing aids in noisy environments. For example, Marrone *et al.* (2008a) showed that acoustic hearing aids did not greatly improve the ability to benefit from spatial separation between a target talker and interfering talkers. They reported a significant but small speech intelligibility benefit from bilateral hearing aids over no hearing aids when three talkers were separated (-90°, 0°, and +90°) than when they were collocated (0°). Even with hearing aids, hearing-impaired listeners showed substantially less improvement in the accuracy of speech intelligibility from spatial separation than normally-hearing listeners. The results of Marrone *et al.* are consistent with the current results showing a small but significant improvement in the accuracy of speech intelligibility with hearing aids but similar EEG responses between trials in which a visual cue had implications for auditory attention and trials in which the cues had no such implications.

The current results did find a significant main effect of hearing aids on amplitudes for some of the Type-I clusters that occurred during the Selective Phase (Table C.1). This result shows that hearing aids modulated the overall responses to the acoustical stimuli at some points during the task. This finding is consistent with previous results showing an effect of hearing aids on EEG responses to acoustical stimuli. For example, Korczak, Kurtzberg, and Stapells (2005) measured ERPs while hearing-impaired adults listened to speech syllables passively and during active discrimination of the syllables /ba/ and /da/. They analysed the amplitudes of the N1, N2, and P3 components. The results showed higher amplitudes during the aided compared to the unaided condition. This result demonstrates that hearing aids have the potential to influence the amplitude of early (< 150 ms) and later (~ 300 ms) EEG responses following the onset of short acoustical stimuli (with 150 ms duration). The current experiment showed, overall, more positive amplitudes with hearing aids in anterior electrodes and more negative amplitudes with hearing aids at posterior electrodes during the acoustical stimuli (Figure C.3E-F and M-N and Figure C.2). However, hearing aids did not significantly affect responses during the Preparatory Phase (Table C.1).

One possible limitation of these analyses is that comparisons between the subset of children who performed the aided and unaided sessions might have been underpowered by the small sample size. As a consequence, the analyses might underestimate the degree to which ERPs differed between sessions. Nevertheless, there were two key aspects of the results which implied that aiding did not significantly modulate differential responses between the Test and Control Conditions. First, all of the *p*-values for the interactions between aiding and condition were greater than 0.1 (and in all but one instance greater than 0.3; Table C.1), which suggested that there was no trend towards a significant interaction. Second, the pattern of amplitudes in the clusters displayed in Figures C.3 and C.4 were similar in the aided and unaided sessions. Therefore, the absence of significant interactions is likely to reflect similarities in ERPs between sessions, rather than low statistical power for detecting significant differences.

C.3.1 Conclusions

Overall, similar patterns of ERPs were recorded when hearing-impaired participants performed a three-talker listening task with and without their acoustic hearing aids. Therefore, the results do not provide evidence that lack of audibility or unfamiliar listening situations underlie the differences between normally-hearing and hearing-impaired children reported in Experiment 4 of Chapter 6.

Appendix D

Comparison of Correct and Incorrect Trials in ERP Analyses

Experiment 4 reported in Chapter 6 included correct and incorrect trials in the analyses for hearing-impaired children. The rationale was that hearing-impaired participants performed with low accuracy and, therefore, removing all incorrect trials would lead to lower signal-to-noise ratio (SNR) in the average event-related potentials (ERPs) for individual participants. Consequently, there would be lower statistical power available to detect differences between normally-hearing and hearing-impaired children during correct-only trials than correct-and-incorrect trials. However, one limitation of including incorrect trials in the analysis is that differences between normally-hearing children have the potential to reflect differences in behavioural performance, rather than the EEG activity that accompanied successful trials (which might produce confounds, for example, if one group was not engaged in the task for all trials of the experiment[12]). This appendix reports analyses for ERPs during correct-only trials in hearing-impaired children, with the aim of identifying possible differences between correct-only and correct-and-incorrect trials.

---

[12] To expand upon this point, I will describe a hypothetical situation, in which ERPs are measured in two different groups of participants. In this hypothetical situation, both groups are capable of performing the task with high accuracy and evoke identical brain activity when they perform the task, but one group is not engaged in the experimental task for all trials of an experiment. If correct and incorrect trials are included in the analysis of ERPs, then there is the potential for a spurious effect for the amplitude of ERPs between different populations. Rather than differences in the ERPs evoked when participants perform the task successfully, differences would result from disengagement in the task. Disengagement would manifest as poorer accuracy and atypical ERPs, even though both groups displayed similar brain activity during successful trials. It is, therefore, important to rule out the explanation that differences in engagement with the task contribute to differences in ERPs between two groups of participants with different levels of behavioural performance.

D.1. Methods

D.1.1 Participants

Participants were 13 children (3 male), aged 7–16 years (mean [M] = 11.9, standard deviation [SD] = 3.0) who completed the first session of Experiment 4 reported in Chapter 6. Details of these participants are reported in Section 6.6.1.1.

D.1.2 Apparatus, Stimuli, and Procedure

The apparatus, stimuli, and procedure are reported in Experiment 4 of Chapter 6.

D.1.3 EEG Recording and Preprocessing

EEG recording and processing were identical to Experiment 3 reported in Chapter 6: Incorrect trials were excluded from the analyses.

D.1.4 Analyses of ERPs

D.1.4.1 Spatio-temporal Cluster-based Permutation Analyses

As a first step, a Spatio-temporal Cluster-based Permutation Analysis was performed. Type-I analyses compared the Test and Control conditions, separately for Location and Gender trials. Type-II analyses compared Location and Gender trials in the Test Condition.

D.1.4.2 Comparisons between correct-only and correct-and-incorrect trials

To investigate whether significantly different patterns of amplitudes occurred for correct-only trials than correct-and-incorrect trials, the average amplitude of activity in each cluster from correct-and-incorrect trials (reported in Experiment 4 of Chapter 6)—averaged over the space-by-time points in the cluster—was compared directly between correct-only trials and correct-and-incorrect trials.

A second analysis compared average amplitudes at the space-by-time points of significant clusters that were identified from correct-only trials in normally-hearing children (reported in Experiment 3 of Chapter 6). This analysis aimed to rule out the explanation that differences between normally-hearing and hearing-impaired children reported in Chapter 6 could be explained by including incorrect trials for hearing-impaired children. The average amplitude of activity in each cluster from normally-hearing children—averaged over the space-by-time points in the cluster—was compared directly in hearing-impaired children between correct-only trials and correct-and-incorrect trials.

D.2. Results

D.2.1.1 Spatio-temporal Cluster-based Permutation Analyses

Type-I analyses: Differences between Test and Control Conditions

Location trials

Figure D.1 illustrates the results of the Type-I analyses on trials in which a Location cue (left/right) was presented. During the 2000-ms Preparatory Phase, no significant clusters of activity were identified. During the Selective Phase, one significant cluster of activity was identified (Figure D.1A). Cluster 18 (Figure D.1B) involved 23 central and posterior electrodes and spanned the time interval from 15 to 463 ms, relative to the start of the phase. Cluster 18 showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 10491, p = 0.019]. The polarity, location, onset time, and duration of Cluster 1 are tabulated in the first line of the second column of Table D.1.

Gender trials

In the second of the Type-I analyses, ERPs between the Test and Control Conditions were compared on trials in which a Gender cue (male/female) was presented. Panels C–E of Figure D.1 illustrate these results. During the 2000-ms Preparatory Phase, no significant clusters of activity were identified. During the Selective Phase, two significant clusters were identified towards the end of the phase (Figure D.1C). Cluster 19 (Figure D.1D) involved 25 central and posterior electrodes and spanned the time interval from 878 to 1200 ms relative to the start of the phase. It showed significantly more positive amplitude during the Test Condition than the Control Condition [cluster statistic = 10205, p = 0.021]. Cluster 20 (Figure D.1E; 912 to 1200 ms) was complementary to Cluster 19. Cluster 20 involved 18 mainly anterior electrodes and showed significantly more negative amplitude during the Test Condition than the Control Condition [cluster statistic = 10304, p = 0.017].

Type-II analyses: Differences between Location and Gender trials

The Type-II analyses did not identify any significant clusters that differed in the Test Condition between Location and Gender trials, either during the Preparatory or Selective Phase.

D.2.1.2 Comparisons between correct-only and correct-and-incorrect trials

Comparisons for clusters identified in correct-and-incorrect trials for hearing-impaired children

To establish whether similar activity arose in correct-only trials and in correct-and-incorrect trials, average amplitudes were compared within each of the Type-I clusters identified in hearing-impaired children for correct-and-incorrect trials (reported in Experiment 4 of Chapter 6). Figure D.2 illustrates the results of these comparisons. The

Location trials

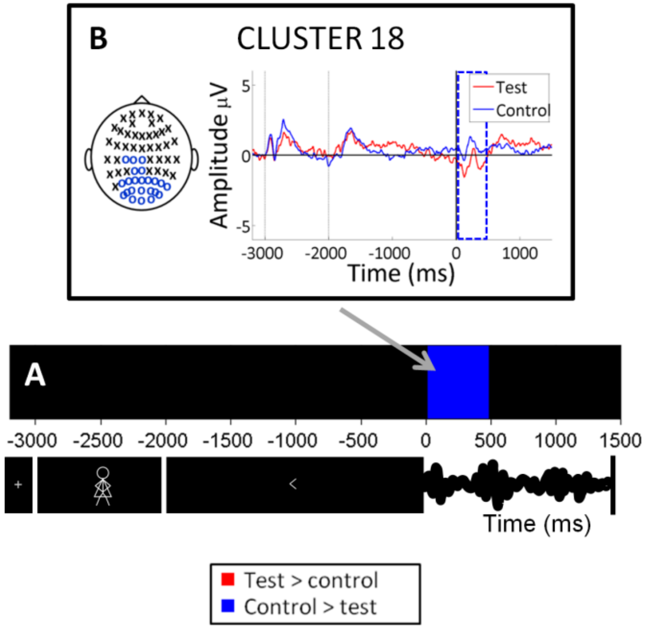


Figure D.1. (Continued on next page). Experiment 4 correct-only trials: Results from Type-I Spatio-temporal Cluster-based Permutation Analyses for Location (**A** to **B**) and Gender (**C** to **E**) trials. (**A** and **C**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition. Further information about each cluster is displayed in (**B**) and (**D**)-(**E**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster, the graph shows the ERPs averaged across those electrodes over the time course of the trial, and the time-span of the cluster is indicated by a dashed rectangle.

Gender trials

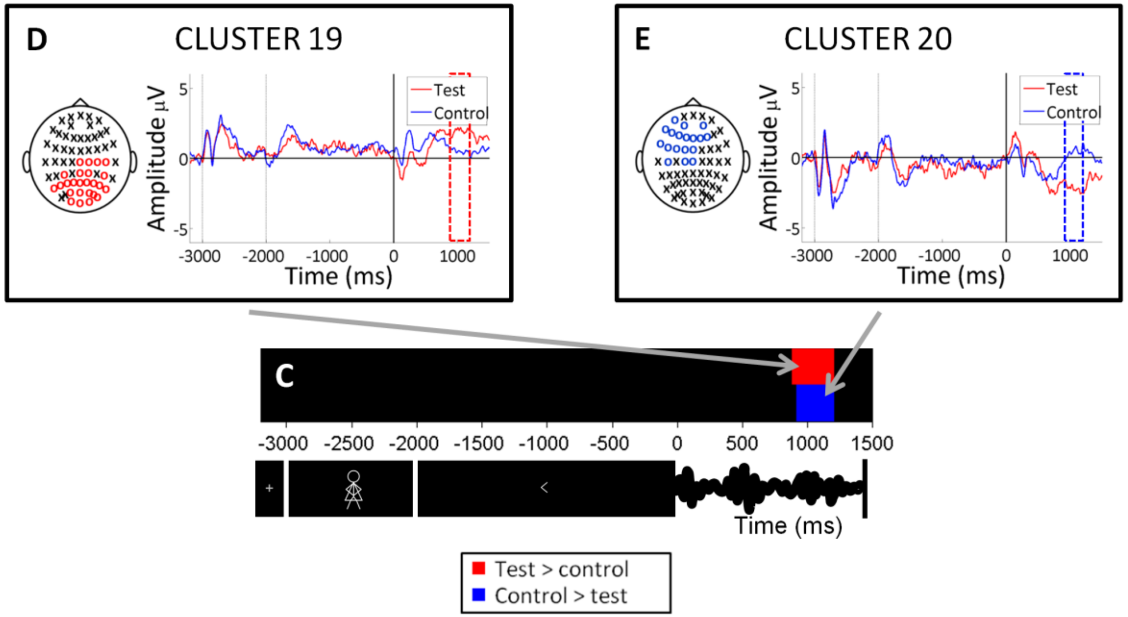


Figure D.1. (Continued from the previous page)

Table D.1. Summary of results for the Gender and Location Condition comparisons between the Test and Control Conditions (Type-I analysis). The results from correct-and-incorrect trials, which are reported in Experiment 4 of Chapter 6, are displayed in the third and fifth columns as a comparison.

Phase		**Correct-and-incorrect Location**	**Correct-only Location**	**Correct-and-incorrect Gender**	**Correct-only Gender**
Preparatory	**Cluster Number**	**-**	**-**	**15**	**-**
	Cluster *p*-value	-	-	0.029	-
	Polarity	-	-	Control > Test	-
	Electrode Locations	-	-	Central + Posterior	-
	Onset of cluster (ms)	-	-	0	-
	Duration of cluster (ms)	-	-	452	-
Selective	**Cluster Number**	**14**	**18**	**16**	**-**
	Cluster *p*-value	0.020	0.019	0.001	-
	Polarity	Control > Test	Control > Test	Control > Test	-
	Electrode Locations	Central + Posterior	Central + Posterior	Central + Posterior	-
	Onset of cluster (ms)	12	15	12	-
	Duration of cluster (ms)	469	448	531	-
Selective	**Cluster Number**	**-**	**-**		**19**
	Cluster *p*-value	-	-		0.021
	Polarity	-	-		Test > Control
	Electrode Locations	-	-		Central + Posterior
	Onset of cluster (ms)	-	-		878
	Duration of cluster (ms)	-	-		322
Selective	**Cluster Number**	**-**	**-**	**17**	**20**
	Cluster *p*-value	-	-	0.033	0.017
	Polarity	-	-	Control > Test	Control > Test
	Electrode Locations	-	-	Anterior	Anterior
	Onset of cluster (ms)	-	-	910	912
	Duration of cluster (ms)	-	-	290	288

difference between the Test and Control Conditions was significant for correct-only trials in all of the clusters identified in correct-and-incorrect trials. The *p*-values that arose from a paired-samples *t*-test (in which average amplitudes in correct-only trials were compared between the Test and Control Conditions) are tabulated in the fourth column of Table D.2.

A 2 x 2 within-subjects ANOVA with the factors trial type (correct-and-incorrect/correct-only) and condition (Test/Control) showed no significant main effect of trial type and no significant interaction between trial type and condition in any of the clusters. The *p*-values are tabulated in the fifth and sixth columns of Table D.2.

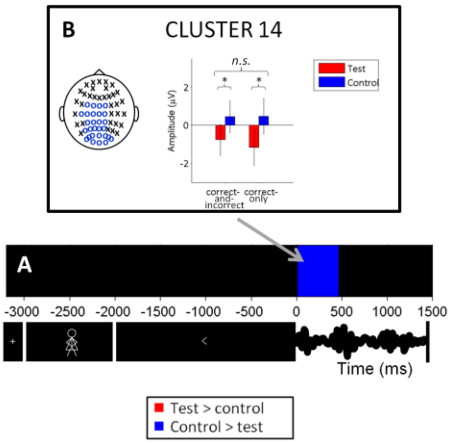
Comparisons for clusters identified in normally-hearing children

To identify whether the comparisons between normally-hearing children and hearing-impaired children reported in Experiment 4 of Chapter 6 reflected activity that was specific to correct-and-incorrect trials or activity that was also similar on correct-only trials, average amplitudes between correct-and-incorrect and correct-only trials were compared

Table D.2. Summary of within-subjects comparisons for hearing-impaired children (n = 13) between cases in which correct and incorrect trials were included in the average waveforms (‘correct-and-incorrect’ trials) and in which only correct trials were included (‘correct-only’ trials). Amplitudes—averaged over the electrodes and time points that contribute to each cluster—were analysed for the Type-I clusters identified for correct-and-incorrect trials from hearing-impaired children in Experiment 4. A tick in the column headed ‘Significant difference between Test/Control Conditions?’ indicates that a paired-samples *t*-test revealed a significant difference in amplitude between the Test and Control Conditions for correct-only trials (*p*-values displayed underneath). The two columns on the far right show the results of a 2 x 2 within-subjects ANOVA with the factors trial type (correct-and-incorrect/correct-only) and condition (Test/Control). A tick in the column headed ‘Main effect of trial type?’ indicates a significant difference in average amplitudes between correct-and-incorrect trials and correct-only trials, and a tick in the column headed ‘Significant interaction between trial types?’ indicates that the ANOVA revealed a significant two-way interaction.

Cue Type	**Phase**	**Cluster Number**	**Significant difference between Test/Control Conditions?**	**Main effect of trial type?**	**Significant interaction between trial types?**
Location	Selective	**14**	✓ $p = 0.018$	✗ $p = 0.37$	✗ $p = 0.25$
Gender	Preparatory	**15**	✓ $p = 0.031$	✗ $p = 0.38$	✗ $p = 0.17$
	Selective	**16**	✓ $p < 0.001$	✗ $p = 0.52$	✗ $p = 0.52$
		17	✓ $p = 0.001$	✗ $p = 0.60$	✗ $p = 0.43$

Location trials



Gender trials

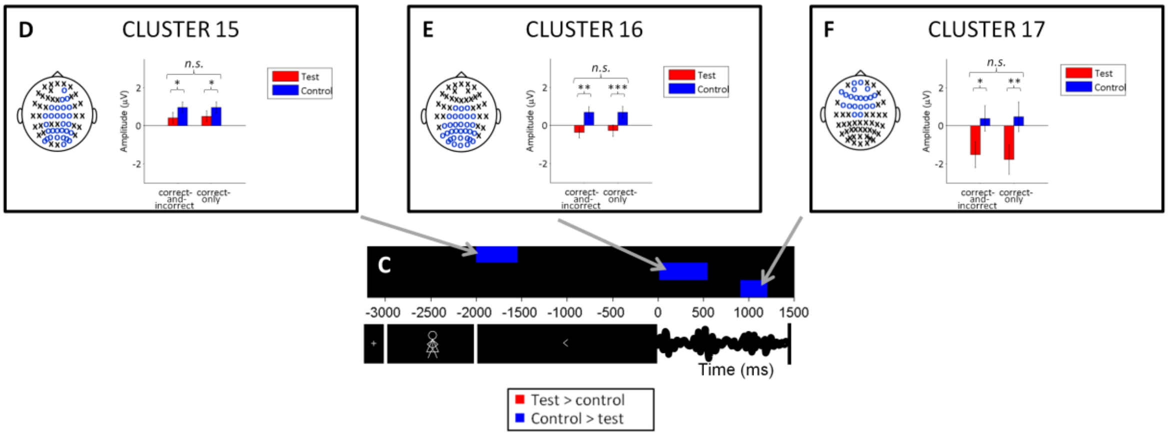


Figure D.2. Comparison of amplitudes in each cluster identified in the correct-and-incorrect Type-I Spatio-temporal Cluster-based Permutation Analyses of Experiment 4 for Location (**A** to **B**) and Gender trials (**C** to **F**) between correct-and-incorrect and correct-only analyses. (**A** and **C**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from correct-and-incorrect analyses. Time on the x-axis is relative to the onset of the acoustical stimuli. Rows on the y-axis show separate significant clusters from correct-and-incorrect analyses. For clusters plotted as red rectangles, the average amplitude, over all space-by-time points in the cluster, was more positive in the Test Condition than the Control Condition in correct-and-incorrect analyses. For clusters plotted as blue rectangles, the average amplitude was more negative in the Test Condition than the Control Condition in correct-and-incorrect analyses. Further information about each cluster is displayed in (B) and (**D**)-(**F**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster in correct-and-incorrect analyses, and the bar graph shows the average amplitude of ERPs—averaged over the electrodes and time points that contributed to the cluster—for correct-and-incorrect analyses and correct-only analyses. Error bars show within-subjects 95% confidence intervals. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing the Test and Control Conditions within each hearing group (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). The larger brackets at the top of each graph indicate the significance level of the interaction between trial type (correct-and-incorrect/correct-only) and the Test/Control Conditions.

Table D.3. Summary of within-subjects comparisons for hearing-impaired children (n = 13) between cases in which correct and incorrect trials were included in the average waveforms ('correct-and-incorrect' trials) and in which only correct trials were included ('correct-only' trials). Amplitudes—averaged over the electrodes and time points that contribute to each cluster—were analysed for the Type-I clusters identified from normally-hearing children in Experiment 3. The column headed 'Significant difference between Test/Control Conditions (correct-and-incorrect)?' shows the previous results for correct-and-incorrect trials reported in Table 6.4: a tick indicates that a paired-samples *t*-test revealed a significant difference in amplitude between the Test and Control Conditions for correct-only trials (*p*-values displayed underneath). The column headed 'Significant difference between Test/Control Conditions (correct-only)?' shows the same analysis performed on correct-only trials. The two columns on the far right show the results of a 2 x 2 within-subjects ANOVA with the factors trial type (correct-and-incorrect/correct-only) and condition (Test/Control). A tick in the column headed 'Main effect of trial type?' indicates a significant difference in average amplitudes between correct-and-incorrect trials and correct-only trials, and a tick in the column headed 'Significant interaction between trial types?' indicates that the ANOVA revealed a significant two-way interaction.

Cue Type	Phase	Cluster Number	Significant difference between Test/Control Conditions (correct-and-incorrect)?	Significant difference between Test/Control Conditions (correct-only)?	Main effect of trial type?	Significant interaction between trial types?
Location	Preparatory	**1**	✗ $p = 0.66$	✗ $p = 0.95$	✗ $p = 0.78$	✗ $p = 0.26$
		2	✗ $p = 0.69$	✗ $p = 0.80$	✗ $p = 0.66$	✗ $p = 0.84$
		3	✗ $p = 0.64$	✗ $p = 0.53$	✗ $p = 0.49$	✗ $p = 0.71$
	Selective	**4**	✓ $p = 0.040$	✓ $p = 0.016$	✗ $p = 0.38$	✗ $p = 0.30$
		5	✗ $p = 0.09$	✗ $p = 0.58$	✗ $p = 0.99$	✗ $p = 0.30$
		6	✗ $p = 0.17$	✗ $p = 0.63$	✗ $p = 0.06$	✗ $p = 0.07$
Gender	Preparatory	**7**	✗ $p = 0.38$	✗ $p = 0.39$	✗ $p = 0.61$	✗ $p = 0.62$
	Selective	**8**	✓ $p < 0.001$	✓ $p = 0.008$	✗ $p = 0.78$	✗ $p = 0.76$
		9	✓ $p = 0.001$	✓ $p = 0.010$	✗ $p = 0.74$	✗ $p = 0.64$
		10	✗ $p = 0.33$	✗ $p = 0.21$	✗ $p = 0.21$	✗ $p = 0.31$
		11	✗ $p = 0.39$	✗ $p = 0.18$	✗ $p = 0.16$	✗ $p = 0.19$

Location trials

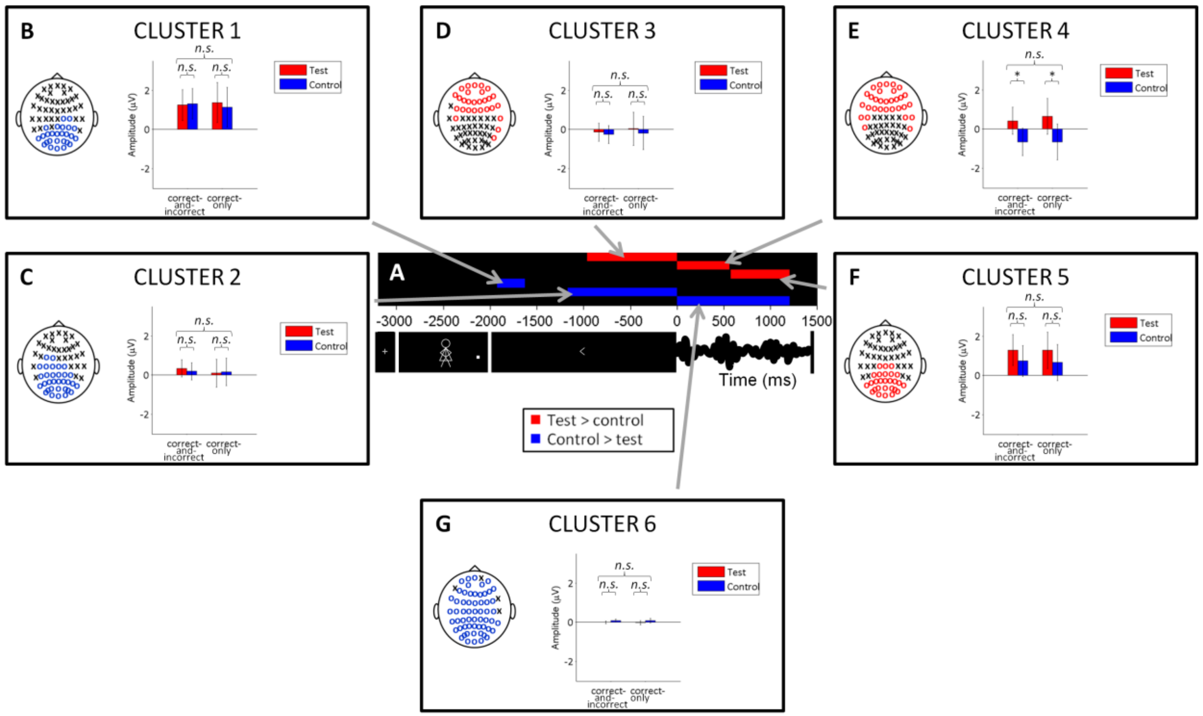


Figure D.3. (Continued on next page). Comparison of amplitudes in each cluster identified in the Type-I Spatio-temporal Cluster-based Permutation Analyses of Experiment 3 for Location trials (**A** to **G**) and Gender trials (**J** to **N**) between correct-and-incorrect and correct-only analyses for hearing-impaired children. (**A** and **H**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. Further information about each cluster is displayed in (**B**)-(**G**) and (**J**)-(**N**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the bar graph shows the average amplitude of ERPs—averaged over the electrodes and time points that contributed to the cluster— for correct-and-incorrect analyses and correct-only analyses. Error bars show within-subjects 95% confidence intervals. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing the Test and Control Conditions within each hearing group (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). The larger brackets at the top of each graph indicate the significance level of the interaction between trial type (correct-and-incorrect/correct-only) and condition (Test/Control).

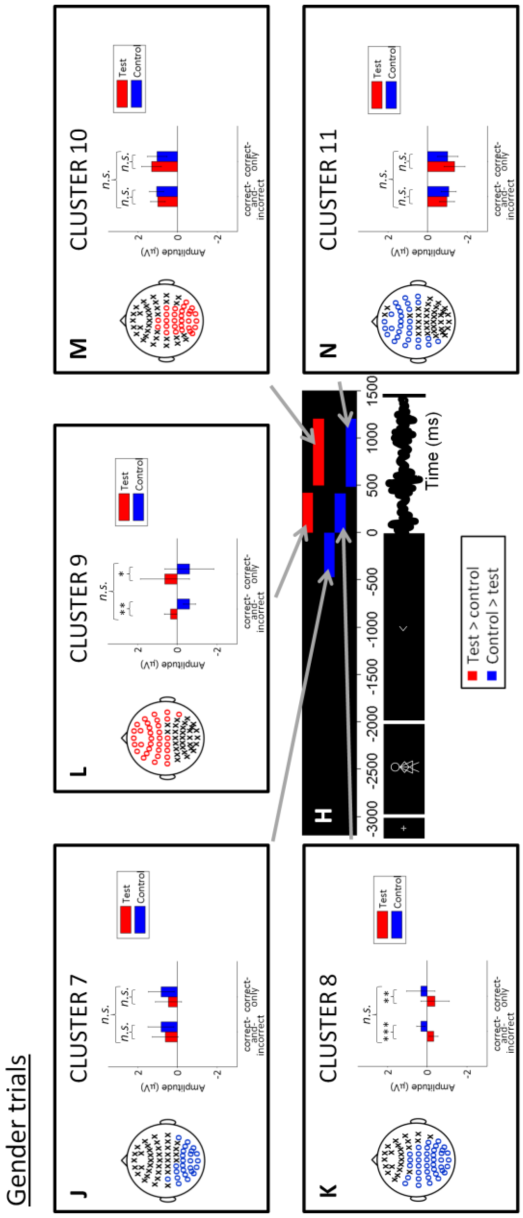


Figure D.3. (Continued from the previous page) within the Type-I and Type-II clusters identified in normally-hearing children (reported in Experiment 3 of Chapter 6). Figures D.3 and D.4 illustrate the results of these comparisons.

A paired-sampled *t*-test compared the average amplitude in correct-only trials between the Test and Control Conditions for each cluster. The resulting *p*-values for the Type-I clusters are tabulated in the fifth column of Table D.3 and for the Type-II clusters in the fifth column of Table D.4. As a comparison, the results reported in Experiment 4 of Chapter 5 for correct-and-incorrect trials are tabulated in the fourth columns of Tables D.3 and D.4. The patterns of significance across the clusters are identical in correct-only and correct-and-incorrect trials (i.e. significant differences either occurred in both correct-only and correct-and-incorrect analyses or in neither of the analyses).

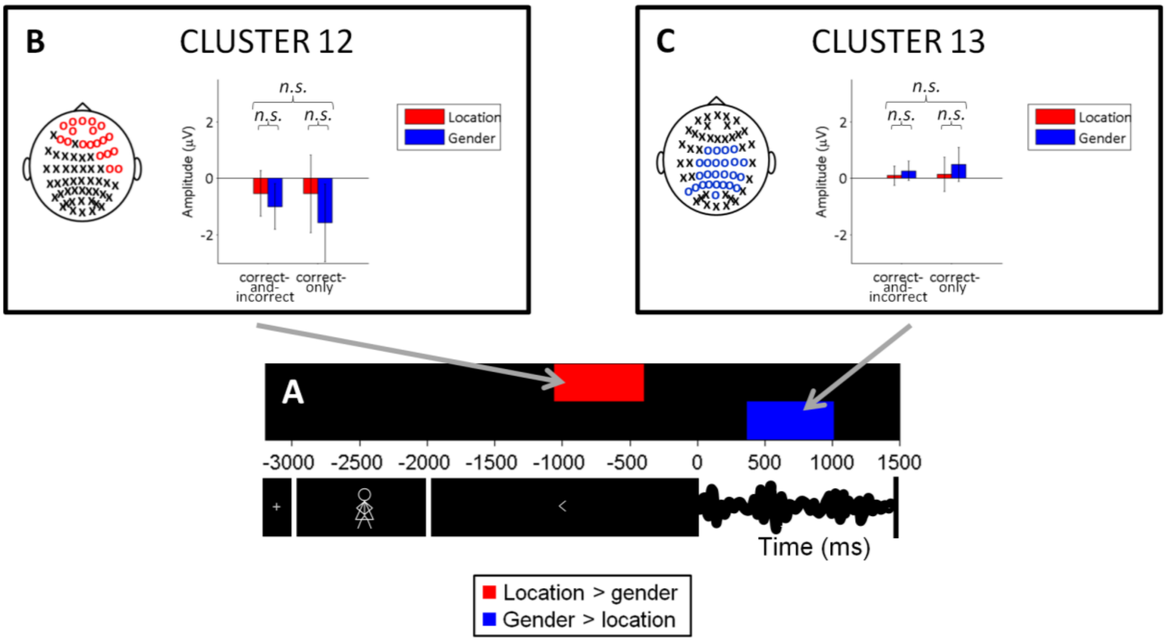


Figure D.4. Comparison of amplitudes in each cluster identified in the Type-II Spatio-temporal Cluster-based Permutation Analyses of Experiment 3, which contrasted Location and Gender trials in the Test Condition, between correct-and-incorrect and correct-only analyses for hearing-impaired children. (**A**) Coloured rectangles indicate the time-span of significant ($p < 0.05$) clusters of activity from Experiment 3. Further information about each cluster is displayed in (**B**)-(**C**) where, for each cluster, the topographical map shows the electrodes that contributed to the cluster in Experiment 3, and the bar graph shows the average amplitude of ERPs—averaged over the electrodes and time points that contributed to the cluster— for correct-and-incorrect analyses and correct-only analyses. Error bars show within-subjects 95% confidence intervals. Smaller brackets displayed on the bar graphs indicate the significance level of a paired-samples *t*-test comparing Location and Gender trials within each hearing group (*n.s.* $p \geq 0.050$, * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$). The larger brackets at the top of each graph indicate the significance level of the interaction between trial type (correct-and-incorrect/correct-only) and the cue type (Location/Gender).

Table D.4. Summary of within-subjects comparisons for hearing-impaired children (n = 13) between cases in which correct and incorrect trials were included in the average waveforms ('correct-and-incorrect' trials) and in which only correct trials were included ('correct-only' trials). Amplitudes—averaged over the electrodes and time points that contribute to each cluster—were analysed for the Type-II clusters identified from normally-hearing children in Experiment 3. The column headed 'Significant difference between Location/Gender trials (correct-and-incorrect)?' shows the previous results for correct-and-incorrect trials reported in Table 6.4: a tick indicates that a paired-samples *t*-test revealed a significant difference in amplitude between the Location and Gender trials for the correct-and-incorrect analysis (*p*-values displayed underneath). The column headed 'Significant difference between Location/Gender trials (correct-only)?' shows the same for the correct-only analysis. The two columns on the far right show the results of a 2 x 2 within-subjects ANOVA with the factors trial type (correct-and-incorrect/correct-only) and cue type (Location/Gender). A tick in the column headed 'Main effect of trial type?' indicates a significant difference in average amplitudes between correct-and-incorrect trials and correct-only trials, and a tick in the column headed 'Significant interaction between trial types?' indicates that the ANOVA revealed a significant two-way interaction.

Cue Type	Phase	Cluster Number	Significant difference between Location/Gender trials (correct-and-incorrect)?	Significant difference between Location/ Gender trials (correct-only)?	Main effect of trial type?	Significant interaction between trial types?
Test	Preparatory	**12**	✗ *p* = 0.29	✗ *p* = 0.98	✗ *p* = 0.71	✗ *p* = 0.39
	Selective	**13**	✗ *p* = 0.36	✗ *p* = 0.51	✗ *p* = 0.08	✗ *p* = 0.86

A 2 x 2 within-subjects ANOVA with the factors trial type (correct-and-incorrect/correct-only) and condition (Test/Control) showed no significant main effect of trial type and no significant interaction between trial type and condition in any of the Type-I clusters. The *p*-values are tabulated in the sixth and seventh columns of Table C.3. A 2 x 2 within-subjects ANOVA with the factors trial type (correct-and-incorrect/correct-only) and cue type (Location/Gender) showed no significant main effect of trial type and no significant interaction between trial type and cue type in any of the Type-II clusters (Table C.4).

D.3. Discussion

Overall, correct-only trials showed similar patterns of amplitudes as identified in correct-and-incorrect trials (Figures D.2–4). The Spatio-temporal Cluster-based Permutation Analysis revealed two clusters during correct-only trials that were similar to the clusters identified during correct-and-incorrect trials (Figure D.1)—Clusters 18 and 20 had similar timing, polarity, and scalp distribution to Clusters 14 and 17 that were identified in correct-and-incorrect trials (Table D.1).

One difference was that, in contrast to the analysis of correct-and-incorrect trials, no significant clusters of activity were identified during the Preparatory Phase of the Gender Condition for the correct-only analysis. There are two possible explanations for this result. First, the cluster might not have emerged due to lower statistical power for correct-only waveforms, due to the contribution of fewer trials to the average waveforms in correct-only than correct-and-incorrect analyses. Second, this cluster might reflect an aspect of processing that participants performed more consistently on incorrect than correct trials. For example, it might reflect an aspect of distraction from the task, which undermined accuracy. The results displayed in Table D.2 support the explanation of lower statistical power—for correct-only trials, there was a significant difference between the Test and Control Conditions at the electrodes and time points of the Preparatory cluster for Gender that was identified in correct-and-incorrect trials. This result is consistent with the explanation that correct-only waveforms were characterised by lower SNRs than correct-and-incorrect trials, which led to lower statistical power for detecting differences between the Test and Control Conditions in the Cluster-based Permutation Analysis.

A second difference between correct-only and correct-and-incorrect trials was a different cluster during the Selective Phase of Gender trials (Figures D.1 and 6.15, respectively). In correct-and-incorrect trials, one cluster emerged early after the talkers began (12 ms) and was sustained for more than 500 ms. In correct-only trials, this cluster did not emerge (Table D.1). Again, the results displayed in Table D.2 suggest that the reason the earlier cluster did not emerge in correct-only trials was due to lower statistical power. In addition, a different cluster emerged later during the Selective Phase of correct-only trials. It is possible that the emergence of a cluster in correct-only trials that was not present in correct-and-incorrect trials reflects an aspect of processing that led to higher-amplitude activity on the scalp during correct than incorrect trials. As a result, including incorrect trials may have increased the amount of noise that resulted from brain activity that was not beneficial to accuracy (such as distraction from the task).

Overall, the results provide strong evidence for similar EEG activity during correct-only and correct-and-incorrect trials. The Cluster-based Permutation Analysis, which incorporated no *a priori* assumptions revealed two clusters of activity that had similar timing and scalp distribution in correct-only and correct-and-incorrect analyses. In addition, clusters revealed on correct-and-incorrect trials that were not revealed during correct-only trials were unlikely to reflect differences in processing between correct and incorrect trials because the same pattern of amplitudes were observed during correct-only trials (Table D.2 and Figure D.2).

In addition, a comparison of amplitudes between correct-only and correct-and-incorrect trials for the clusters identified in normally-hearing children (Tables D.3–4) showed similar patterns of amplitudes across correct-only and correct-and-incorrect trials (Figures D.3–4). This result suggests that differences between normally-hearing and hearing-impaired children reported in Experiment 4 of Chapter 6 cannot be explained by the inclusion of incorrect trials for hearing-impaired children. The results provide strong evidence for similar patterns of amplitudes during correct-only as correct-and-incorrect trials in hearing-impaired children. Therefore, the differences between normally-hearing and hearing-impaired children identified in Experiment 4 of Chapter 6 are likely to reflect differences in brain activity evoked during multi-talker listening, rather than differences in the trials that were included in the analyses.

One possible limitation of the analyses reported in this appendix was that correct-only trials were compared with correct-and-incorrect trials, rather than comparing correct-only and incorrect-only trials directly. The main reason for this decision was that the aim of this appendix was to compare the method employed for the results from hearing-impaired children in Chapter 6 (correct-and-incorrect) with the method employed for the results from normally-hearing children in Chapter 6 (correct-only) to explore whether different results would have been obtained had both analyses used correct-only trials, as would be most desirable. Nevertheless, in future research, it might be useful to correct-only and incorrect-only trials, since this would address the question of whether different EEG activity accompanies trials in which hearing-impaired children are able to correctly identify the colour and number spoken by a target talker in a mixture of talkers and trials in which they are not able to report words spoken by the target talker.

D.3.1 Conclusions

Overall, the results provide strong evidence for similar patterns of amplitudes when only correct trials are included in the analyses for hearing-impaired children as when correct and incorrect trials are included. This finding suggests that differences between normally-hearing and hearing-impaired children reported in Experiment 4 of Chapter 6 reflect differences in the brain activity evoked during *successful* speech intelligibility during multi-talker listening.

Appendix E

RTs of individual participants in Chapter 4

Figure E.1 shows the reaction times (RTs) of individual adult participants in Experiment 2 of Chapter 4.

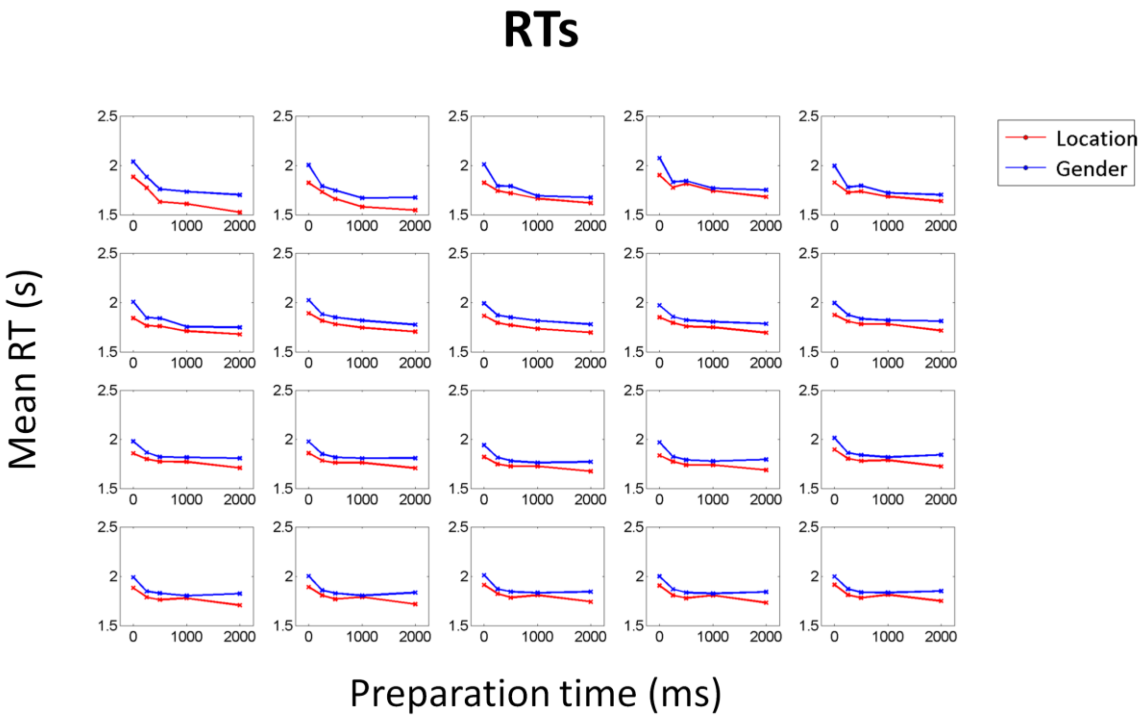


Figure E.1. RTs for individual participants in Experiment 2 of Chapter 4. Each graph displays average RTs of an individual participant for the five different preparation time conditions.

References

Adriani, M., Maeder, P. P., Meuli, R., Thiran, A. B., Frischknecht, R., Villemure, J.-G., … Clarke, S. (2003). Sound recognition and localization in man: Specialized cortical networks and effects of acute circumscribed lesions. *Experimental Brain Research*, *153*(4), 591–604.

Ahmadlou, M., & Adeli, H. (2011). Functional community analysis of brain: a new approach for EEG-based investigation of the brain pathology. *NeuroImage*, *58*(2), 401–8.

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M. S., … Belliveau, J. W. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(39), 14608–13.

Alain, C., & Arnott, S. R. (2000). Selectively attending to auditory objects. *Front. Biosci*, (10), 202–212.

Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(21), 12301–6.

Alain, C., & Tremblay, K. L. (2007). The role of event-related brain potentials in assessing central auditory processing. *Journal of the American Academy of Audiology*, *18*(7), 573–89.

Allen, K., Alais, D., Shinn-Cunningham, B. G., & Carlile, S. (2011). Masker location uncertainty reveals evidence for suppression of maskers in two-talker contexts. *The Journal of the Acoustical Society of America*, *130*(4), 2043–53.

Allen, K., Carlile, S., & Alais, D. (2008). Contributions of talker characteristics and spatial location to auditory streaming. *The Journal of the Acoustical Society of America*, *123*(3), 1562–70.

Allport, D. A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (pp. 183–218). Cambridge, Massechusetts: MIT Press.

Amitay, S., Zhang, Y.-X., Jones, P. R., & Moore, D. (2014). Perceptual learning: top to bottom. *Vision Research*, *99*, 69–77.

Arbogast, T. L., & Kidd, G. (2000). Evidence for spatial tuning in informational masking using the probe-signal method. *The Journal of the Acoustical Society of America*, *108*(4), 1803–10.

Arbogast, T. L., Mason, C. R., & Kidd, G. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *117*(4), 2169–2180.

Arehart, K. H., King, C. A., & McLean-Mudgett, K. S. (1997). Role of Fundamental Frequency Differences in the Perceptual Separation of Competing Vowel Sounds by Listeners With Normal Hearing and Listeners With Hearing Loss. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1434–1444.

Assmann, P. F., & Summerfield, A. Q. (1994). The contribution of waveform interactions to the perception of concurrent vowels. *The Journal of the Acoustical Society of America*, *95*(1), 471–484.

Bachorowski, J. A., & Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America*, *106*(2), 1054–63.

Baer, T., & Moore, B. (1994). Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *The Journal of the Acoustical Society of America*, *95*(4), 2277–80.

Bamiou, D.-E., Musiek, F. E., & Luxon, L. M. (2001). Aetiology and clinical presentations of auditory processing disorders--a review. *Archives of Disease in Childhood*, *85*(5), 361–5.

beim Graben, P. (2001). Estimating and improving the signal-to-noise ratio of time series by symbolic dynamics. *Physical Review E*, *64*(5), 051104.

Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*, *14*(16), 2105–9.

Bellmann, A, Meuli, R., & Clarke, S. (2001). Two types of auditory neglect. *Brain*, *124*, 676–87.

Bertrand, O., Perrin, F., & Pernier, J. (1985). A theoretical justification of the average reference in topographic evoked potential studies. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *62*(6), 462–4.

Best, V., Gallun, F. J., Carlile, S., & Shinn-Cunningham, B. G. (2007). Binaural interference and auditory grouping. *The Journal of the Acoustical Society of America*, *121*(2), 1070.

Best, V., Ihlefeld, A., & Shinn-Cunningham, B. G. (2005). The effect of auditory spatial layout in a divided attention task. *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display*.

Best, V., Marrone, N., Mason, C. R., Kidd, G., & Shinn-Cunningham, B. G. (2009). Effects of sensorineural hearing loss on visually guided attention in a multitalker environment. *Journal of the Association for Research in Otolaryngology*, *10*(1), 142–9.

Best, V., Ozmeral, E. J., & Shinn-Cunningham, B. G. (2007). Visually-guided attention enhances target identification in a complex auditory scene. *Journal of the Association for Research in Otolaryngology*, *8*(2), 294–304.

Bishop, D. V. M., Anderson, M., Reid, C., & Fox, A. M. (2011). Auditory development between 7 and 11 years: an event-related potential (ERP) study. *PloS One*, *6*(5), e18993.

Bonino, A. Y., Leibold, L. J., & Buss, E. (2012). Release from perceptual masking for children and adults: benefit of a carrier phrase. *Ear and Hearing*, *34*(1), 3–14.

Braitenberg, V., & Schüz, A. (1991). *Anatomy of the cortex: statistics and geometry* (p. 249). Springer-Verlag.

Braver, T. S., & West, R. (2008). Working Memory, Executive Control, and Aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (Third Edit.). New York: Psychology Press.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, *13*, 87–100.

Broadbent, D. E. (1958). *Perception and communication* (p. 338). Pergamon Press.

Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica*, *86*, 117–128.

Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of Acoustical Society of America*, *83*(4), 1508–16.

Bronkhorst, A. W., & Plomp, R. (1992). Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *The Journal of the Acoustical Society of America*, 3132–3139.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *109*(3), 1101.

Brungart, D. S., & Simpson, B. D. (2002a). The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal. *The Journal of the Acoustical Society of America*, *112*(2), 664–676.

Brungart, D. S., & Simpson, B. D. (2002b). Within-ear and across-ear interference in a cocktail-party listening task. *The Journal of the Acoustical Society of America*, *112*(6), 2985.

Brungart, D. S., & Simpson, B. D. (2007). Cocktail party listening in a dynamic multitalker environment. *Perception & Psychophysics*, *69*(1), 79–91.

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, *110*(5), 2527–2538.

Burk, M. H., & Humes, L. E. (2007). Effects of Training on Speech Using Lexically Hard Words. *Journal of Speech, Language, and Hearing Research*, *50*, 25–40.

Buzsáki, G., Anastassiou, C. a, & Koch, C. (2012). The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nature Reviews. Neuroscience*, *13*(6), 407–20.

Buzsáki, G., Traub, R., & Pedley, T. (2003). The cellular synaptic generation of EEG activity. In J. S. Ebersole & T. A. Pedley (Eds.), *Current Practice of Clinical Encephalography* (pp. 1–11). Philadelphia: Lippincott Williams & Wilkins.

Caporello Bluvas, E., & Gentner, T. Q. (2013). Attention to natural auditory signals. *Hearing Research*, *305*, 10–8.

Caramazza, a, & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*(1), 1–34.

Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkey, R. H., & Colburn, H. S. (2002). Auditory Phase Opponency: A Temporal Model for Masked Detection at Low Frequencies. *Acta Acustica*, *88*, 334–347.

Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, *2*(7), 671–6.

Choi, I., Wang, L., Bharadwaj, H., & Shinn-Cunningham, B. G. (2014). Individual differences in attentional modulation of cortical responses correlate with selective attention performance. *Hearing Research*, *314*, 10–19.

Coch, D., & Gullick, M. M. (2011). Event-related potentials and development. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 475–512). Oxford University Press.

Conway, B. R. (2009). Color vision, cones, and color-coding in the cortex. *The Neuroscientist*, *15*(3), 274–90.

Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, *119*(3), 1562–1573.

Cooper, R., Winter, A. L., Crow, H. J., & Walter, W. G. (1965). Comparison of subcortical, cortical and scalp activity using chronically indwelling electrodes in man. *Electroencephalography and Clinical Neurophysiology*, *18*(3), 217–228.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*(3), 292–7.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, *58*(3), 306–24.

Corbetta, M., Tansy, A. P., Stanley, C. M., Astafiev, S. V, Snyder, A. Z., & Shulman, G. L. (2005). A functional MRI study of preparatory signals for spatial location and objects. *Neuropsychologia*, *43*(14), 2041–56.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: The neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience*, *18*(18), 7426–35.

Cuffin, B. N. (1993). Effects of Local Variations In Skull and Scalp Thickness on EEG's and MEG's, *40*, 42–48.

Cuffin, B. N. (1996). EEG localization accuracy improvements using realistically shaped head models. *IEEE Transactions on Bio-Medical Engineering*, *43*(3), 299–303.

Darwin, C. J. (2006). Contributions of binaural information to the separation of different sound sources. *International Journal of Audiology*, *45*, S20–4.

Darwin, C. J. (2008). Listening to speech in the presence of other sounds. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1493), 1011–21.

Darwin, C. J., Brungart, D. S., & Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *114*(5), 2913–2922.

Darwin, C. J., & Hukin, R. W. (1999). Auditory objects of attention: the role of interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(3), 617–29.

Darwin, C. J., & Hukin, R. W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *Journal of Acoustical Society of America*, *107*(2), 970–977.

Denys, K., Vanduffel, W., Fize, D., Nelissen, K., Peuskens, H., Van Essen, D., & Orban, G. A. (2004). The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. *The Journal of Neuroscience*, *24*(10), 2551–65.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

Desmedt, J. E., Chalklin, V., & Tomberg, C. (1990). Emulation of somatosensory evoked potential (SEP) components with the 3-shell head model and the problem of "ghost potential fields" when using an average reference in brain mapping. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *77*(4), 243–258.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some Theoretical considerations. *Psychological Review*, *70*(1), 80–90.

Dhamani, I., Leung, J., Carlile, S., & Sharma, M. (2013). Switch attention to listen. *Scientific Reports*, *3*, 1297.

Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. *Behavior Research Methods, Instruments, & Computers*, *30*(1), 34–43.

Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America*, *2012*, 5–10.

Downing, P. E., Chan, A. W.-Y., Peelen, M. V, Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cerebral Cortex*, *16*(10), 1453–61.

Dreschler, W. A., & Plomp, R. (1985). Relations between psychophysical data and speech perception for hearing-impaired subjects. II. *The Journal of the Acoustical Society of America*, *78*(4), 1261–70.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*, 53–78.

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of Acoustical Society of America*, *76*(1), 87–96.

Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *The Quarterly Journal of Experimental Psychology*, *59*(1), 2–27.

Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *The Journal of the Acoustical Society of America*, *74*(3), 739–43.

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America*, *114*(1), 368.

Ebersole, J. S. (1997). Defining epileptogenic foci: past, present, future. *Journal of Clinical Neurophysiology*, *14*(6), 470–83.

Edwards, B. (2007). The future of hearing aid technology. *Trends in Amplification*, *11*(1), 31–45.

Elhilali, M., Fritz, J. B., Chi, T.-S., & Shamma, S. a. (2007). Auditory cortical receptive fields: stable entities with plastic abilities. *The Journal of Neuroscience*, *27*(39), 10372–82.

Ericson, M. A., Brungart, D. S., & Brian, D. (2004). Factors That Influence Intelligibility in Multitalker Speech Displays. *The International Journal of Aviation Psychology*, *14*(3), 313–334.

Falkenberg, L. E., Specht, K., & Westerhausen, R. (2011). Attention and cognitive control networks assessed in a dichotic listening fMRI study. *Brain and Cognition*, *76*(2), 276–85.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *Journal of Acoustical Society of America*, *108*(6), 3023–3029.

Festen, J. M., & Plomp, R. (1983). Relations between auditory functions in impaired hearing. *The Journal of the Acoustical Society of America*, *73*(2), 652–62.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, *88*(4), 1725–36.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America*, *109*(5), 2112–2122.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America*, *115*(5), 2246–2256.

Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America*, *106*(6), 3578–88.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews*, *25*(4), 355–73.

Fries, P. (2001). Modulation of Oscillatory Neuronal Synchronization by Selective Visual Attention. *Science*, *291*(5508), 1560–1563.

Fritz, J. B., Shamma, S. A., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, *6*(11), 1216–23.

Fuchs, M., Wagner, M., & Kastner, J. (2001). Boundary element method volume conductor models for EEG source reconstruction. *Clinical Neurophysiology*, *112*(8), 1400–7.

Gatehouse, S., & Akeroyd, M. A. (2006). Two-eared listening in dynamic situations. *International Journal of Audiology*, *45*, S120–4.

Gatehouse, S., & Akeroyd, M. a. (2008). The effects of cueing temporal and spatial attention on word recognition in a complex listening task in hearing-impaired listeners. *Trends in Amplification*, *12*(2), 145–61.

Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, *43*, 85–99.

Gaudrain, E., Grimault, N., Healy, E. W., & Béra, J. (2007). Effect of spectral smearing on the perceptual segregation of vowel sequences. *Hearing Research*, *231*, 32–41.

Giesbrecht, B., Weissman, D. H., Woldorff, M. G., & Mangun, G. R. (2006). Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*, *1080*(1), 63–72.

Giesbrecht, B., Woldorff, M. G., Song, A. W., & Mangun, G. R. (2003). Neural mechanisms of top-down control during spatial and feature attention. *NeuroImage*, *19*(3), 496–512.

Green, J. J., & McDonald, J. J. (2008). Electrical Neuroimaging Reveals Timing of Attentional Control Activity in Human Brain. *PLoS Biology*, *6*(4), e81.

Greenberg, S., & Ainsworth, W. A. (2006). *Listening to Speech: An Auditory Perspective* (p. 458). Mahwah, NJ: Lawrence Erlbaum.

Gregan, M. J., Nelson, P. B., & Oxenham, A. J. (2013). Behavioral measures of cochlear compression and temporal resolution as predictors of speech masking release in hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *134*(4), 2895–912.

Gregoriou, G. G., Gotts, S. J., Zhou, H., & Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, *324*(5931), 1207–10.

Gregoriou, G. G., Rossi, A. F., Ungerleider, L. G., & Desimone, R. (2014). Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. *Nature Neuroscience*, *17*(7), 1003–11.

Grose, J. H., & Hall, J. W. (1996). Perceptual organization of sequential stimuli in listeners with cochlear hearing loss. *Journal of Speech and Hearing Research*, *39*(6), 1149–58.

Gross, C. G. (1992). Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *335*(1273), 3–10.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol*, *35*(1), 96–111.

Halliday, L. F., & Bishop, D. V. M. (2005). Frequency discrimination and literacy skills in children with mild to moderate sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *48*, 1187–1204.

Hämäläinen, M. S., & Ilmoniemi, R. J. (1984). Interpreting measured magnetic fields of the brain: estimates of current distributions. In *Technical Report TKK-F-A599*. Espoo: Helsinki University of Technology.

Handy, T. C. (2009). *Brain Signal Analysis: Advances in Neuroelectric and Neuromagnetic Methods* (p. 272). Cambridge, Massechusetts: The MIT Press.

Harkrider, A. W., Plyler, P. N., & Hedrick, M. S. (2006). Effects of hearing loss and spectral shaping on identification and neural response patterns of stop-consonant stimuli. *The Journal of the Acoustical Society of America*, *120*(2), 915.

Hawkins, D. B., & Wightman, F. L. (1980). Interaural time discrimination ability of listeners with sensorineural hearing loss. *Audiology*, *19*, 495–507.

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *The Journal of the Acoustical Society of America*, *115*(2), 833–843.

Hegerl, U., Gallinat, J., & Mrowinski, D. (1994). Intensity dependence of auditory evoked dipole source activity. *International Journal of Psychophysiology*, *17*(1), 1–13.

Helfer, K. S., & Freyman, R. L. (2005). The role of visual speech cues in reducing energetic and informational masking. *The Journal of the Acoustical Society of America*, *117*(2), 842–849.

Helfer, K. S., & Freyman, R. L. (2008). Aging and speech-on-speech masking. *Ear and Hearing*, *29*(1), 87–98.

Helmholtz, H. (1853). Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche. *Annalen Der Physik, 89,* 211–33.

Hilbert, D. (1912). *Grundzüge einer Allgemeinen Theorie der linearen Integralgleichungen (Basics of a general theory of linear integral equations)* (pp. 1–310). Leizpig: Teubner.

Hill, K. T., & Miller, L. M. (2010). Auditory attentional control and selection during cocktail party listening. *Cerebral Cortex*, *20*(3), 583–90.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–80.

Hink, R. F., Van Voorhis, S. T., Hillyard, S. A., & Smith, T. S. (1977). The division of attention and the human auditory evoked potential. *Neuropsychologia*, *15*(4-5), 597–605.

Hoffmann, S., & Falkenstein, M. (2008). The correction of eye blink artefacts in the EEG: a comparison of two prominent methods. *PloS One*, *3*(8), e3004.

Horvath, J., Sussman, E., Winkler, I., & Schröger, E. (2011). Preventing distraction: Assessing stimulus-specific and general effects of the predictive cueing of deviant auditory events. *Biological Psychology*, *87*(1), 35–48.

Huppert, F., & Wilcock, G. (1997). Ageing, cognition and dementia. *Age and Ageing*, 20–23.

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, *13*(4-5), 411–30.

Ihlefeld, A., Shinn-Cunningham, B. G., & Carlyon, R. P. (2012). Comodulation masking release in speech identification with real and simulated cochlear-implant hearing. *The Journal of the Acoustical Society of America*, *131*(2), 1315–24.

Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America*, *121*(4), 2283.

Johnsrude, I. S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H. P., & Carlyon, R. P. (2013). Swinging at a cocktail party: voice familiarity aids speech perception in the presence of a competing voice. *Psychological Science*, *24*(10), 1995–2004.

Jonides, J. (1981). Voluntary versus automatic control over the mind’s eye's movement. *Attention and Performance IX*, 187–203.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(2), 163–78.

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*(10), 1745–58.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, *47*(5), 888–904.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*(4), 751–61.

Kawashima, R., Imaizumi, S., Mori, K., Okada, K., Goto, R., Kiritani, S., … Fukuda, H. (1999). Selective visual and auditory attention toward utterances-a PET study. *NeuroImage*, *10*(2), 209–15.

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., … Yee, C. M. (2013). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*.

Keil, J., Müller, N., Ihssen, N., & Weisz, N. (2012). On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cerebral Cortex*, *22*(1), 221–31.

Kidd, G., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listen. *The Journal of the Acoustical Society of America*, *118*(6), 3804.

Kitterick, P. T., Bailey, P. J., & Summerfield, A. Q. (2010). Benefits of knowing who, where, and when in multi-talker listening. *The Journal of the Acoustical Society of America*, *127*(4), 2498–508.

Koch, I., Lawo, V., Fels, J., & Vorländer, M. (2011). Switching in the cocktail party: exploring intentional control of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(4), 1140–7.

Kochkin, B. S. (2005). Customer satisfaction with hearing instruments in the digital age. *The Hearing Journal, 58*(9).

Koessler, L., Maillard, L., Benhadid, A., Vignal, J.-P., Braun, M., & Vespignani, H. (2007). Spatial localization of EEG electrodes. *Clinical Neurophysiology*, *37*(2), 97–102.

Koravand, A., Jutras, B., & Lassonde, M. (2012). Cortical auditory evoked potentials in children with a hearing loss: a pilot study. *International Journal of Pediatrics*, *2012*, 250254.

Korczak, P. a, Kurtzberg, D., & Stapells, D. R. (2005). Effects of sensorineural hearing loss and personal hearing aids on cortical event-related potential and behavioral measures of speech-sound processing. *Ear and Hearing*, *26*(2), 165–85.

Kramer, A. F., & Madden, D. J. (2008). Attention. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (Third Edit., pp. 189–250). New York: Psychology Press.

Kraus, N., & Anderson, S. (2012). Hearing with Our Brains. *The Hearing Journal*, *65*(9), 48.

Lachaux, J.-P., Fonlupt, P., Kahane, P., Minotti, L., Hoffmann, D., Bertrand, O., & Baciu, M. (2007). Relationship between task-related gamma oscillations and BOLD signal: new insights from combined fMRI and intracranial EEG. *Human Brain Mapping*, *28*(12), 1368–75.

Lacher-Fougère, S., & Demany, L. (2005). Consequences of cochlear damage for the detection of interaural phase differences. *The Journal of the Acoustical Society of America*, *118*(4), 2519.

Lantz, G., Grave de Peralta, R., Spinelli, L., Seeck, M., & Michel, C. M. (2003). Epileptic source localization with high density EEG: how many electrodes are needed? *Clinical Neurophysiology*, *114*(1), 63–9.

Larson, E., & Lee, A. K. C. (2013a). Influence of preparation time and pitch separation in switching of auditory attention between streams. *The Journal of the Acoustical Society of America*, *134*(2), EL165–71.

Larson, E., & Lee, A. K. C. (2013b). The cortical dynamics underlying effective switching of auditory spatial attention. *NeuroImage*, *64*, 365–70.

Lavie, N. (2001). Capacity limits in selective attention: behavioral evidence and implications for neural activity. In J. Braun & C. Koch (Eds.), *Visual attention and cortical circuits* (pp. 49–68). Cambridge, Massachusetts: MIT press.

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–97.

Leavitt, V. M., Molholm, S., Gomez-Ramirez, M., & Foxe, J. J. (2011). “What” and “where” in auditory sensory processing: A high-density electrical mapping study of distinct neural processes underlying sound object recognition and sound localization. *Frontiers in Integrative Neuroscience*, *5*, 1–15.

Lebedev, M. a, Messinger, A., Kralik, J. D., & Wise, S. P. (2004). Representation of attended versus remembered locations in prefrontal cortex. *PLoS Biology*, *2*(11), e365.

Lee, A. K. C., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M. S., & Shinn-Cunningham, B. G. (2013). Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. *Frontiers in Neuroscience*, *6*, 1–9.

Lee, C.-C., & Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, *14*(1), 108–114.

Lee, L. W., & Humes, L. E. (1993). Evaluating a speech-reception threshold model for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *93*(5), 2879–85.

Leek, M. R., & Summers, V. (2001). Pitch strength and pitch dominance of iterated rippled noises in hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *109*(6), 2944–54.

Li, S. C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: from neuromodulation to representation. *Trends in Cognitive Sciences*, *5*(11), 479–486.

Lin, F.-H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, *31*(1), 160–71.

Liu, T. (2003). Cortical Mechanisms of Feature-based Attentional Control. *Cerebral Cortex*, *13*(12), 1334–1343.

Loeb, G., White, M., & Merzenich, M. M. (1983). Spatial cross-correlation. *Biological Cybernetics*, *47*, 149–163.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18866–9.

Lu, Z.-L., Tse, H. C.-H., Dosher, B. A., Lesmes, L. A., Posner, C., & Chu, W. (2009). Intra- and cross-modal cuing of spatial attention: Time courses and mechanisms. *Vision Research*, *49*(10), 1081–96.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, Massechusetts: MIT Press.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, *77*(1), 24–42.

Lundeen, C. (1991). Prevalence of Hearing Impairment Among School Children. *Language, Speech, and Hearing Services in Schools*, *22*(1), 269–271.

Lutfi, R. A., Kistler, D. J., Oh, E. L., Wightman, F. L., & Callahan, M. R. (2003). One factor underlies individual differences in auditory informational masking within and across age groups. *Perception & Psychophysics*, *65*(3), 396–406.

Mackersie, C. L., Dewey, J., & Guthrie, L. A. (2011). Effects of fundamental frequency and vocal-tract length cues on sentence segregation by listeners with hearing loss. *The Journal of the Acoustical Society of America*, *130*(2), 1006–19.

Maddox, R. K., & Shinn-Cunningham, B. G. (2012). Influence of task-relevant and task-irrelevant feature continuity on selective auditory attention. *Journal of the Association for Research in Otolaryngology*, *13*(1), 119–29.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–10.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–90.

Marozzi, M. (2004). Some remarks about the number of permutations one should consider to perform a permutation test. *Statistica*, *64*(1), 193–201.

Marrone, N., Mason, C. R., & Kidd, G. (2008a). Evaluating the benefit of hearing aids in solving the cocktail party problem. *Trends Amplif.*, *12*(4), 300–315.

Marrone, N., Mason, C. R., & Kidd, G. (2008b). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *The Journal of the Acoustical Society of America*, *124*(5), 3064–75.

Marrone, N., Mason, C. R., & Kidd, G. (2008c). Tuning in the spatial dimension: evidence from a masked speech identification task. *The Journal of the Acoustical Society of America*, *124*(2), 1146–58.

McCullouch, D. L. (2013). Visual evoked potentials in infants. In M. de Haan (Ed.), *Infant EEG and Event-Related Potentials* (pp. 39–76). Psychology Press.

McEvoy, L., Picton, T. W., Champagne, S. C., Kellett, A. J. C., & Kelly, J. B. (1990). Human evoked potentials to shifts in the lateralization of a noise. *Audiology*, *29*, 163–180.

Meiran, N., Chorev, Z., & Sapir, a. (2000). Component processes in task switching. *Cognitive Psychology*, *41*(3), 211–53.

Menninghaus, E., Lütkenhöner, B., & Gonzalez, S. L. (1994). Localization of a dipolar source in a skull phantom: realistic versus spherical model. *IEEE Transactions on Bio-Medical Engineering*, *41*(10), 986–9.

Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, *485*(7397), 233–6.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, *115*(10), 2195–222.

Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 66–79.

Monsell, S., & Driver, J. (2000). *Control of cognitive processes: Attention and performance XVIII*. (S. Monsell & J. Driver, Eds.). Cambridge, Massechusetts: MIT Press.

Moore, B. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues* (Second Edi., p. 344). John Wiley & Sons.

Moore, B. (2008a). Basic auditory processes involved in the analysis of speech sounds. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1493), 947–63.

Moore, B. (2008b). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology*, *9*(4), 399–406.

Moore, B., & Glasberg, B. R. (1987). Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes. *The Journal of the Acoustical Society of America*, *82*(1), 69–79.

Moore, B., & Sek, A. (1996). Detection of frequency modulation at low modulation rates: evidence for a mechanism based on phase locking. *The Journal of the Acoustical Society of America*, *100*, 2320–31.

Moore, D. (2014). Sources of pathology underlying listening disorders in children. *International Journal of Psychophysiology, 14,* 174.

Moore, D., Rosen, S., Bamiou, D.-E., Campbell, N. G., & Sirimanna, T. (2013). Evolving concepts of developmental auditory processing disorder (APD): A British Society of Audiology APD Special Interest Group “white paper”. *International Journal of Audiology*, *52*(1), 3–13.

Münte, T. F., Spring, D. K., Szycik, G. R., & Noesselt, T. (2010). Electrophysiological attention effects in a virtual cocktail-party setting. *Brain Research*, *1307*, 78–88.

Murry, T., & Singh, S. (1980). Multidimensional analysis of male and female voices. *The Journal of the Acoustical Society of America*, *68*(5), 1294–300.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*, 201–288.

Naatanen, R., Gaillard, A. W. K., & Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313–329.

Nakai, T., Kato, C., & Matsuo, K. (2005). An fMRI study to investigate auditory attention: A model of the cocktail party phenomenon. *Magnetic Resonance in Medical Sciences*, *4*(2), 75–82.

Neher, T., Behrens, T., Carlile, S., Jin, C., Kragelund, L., Petersen, A. S., & Schaik, A. Van. (2009). Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition. *International Journal of Audiology*, *48*(11), 758–74.

Nelson, W. T., Bolia, R. S., Ericson, M. A., & Mckinley, R. L. (1998). Monitoring the simultaneous presentation of spatialized speech signals in a virtual acoustic environment. *Proceedings of the 1998 IMAGE Conference*, 159–166.

Neuper, C., & Klimesch, W. (Eds.). (2006). *Event-Related Dynamics of Brain Oscillations*. Amsterdam: Elsevier.

Niedermeyer, E., & Lopes da Silva, F. H. (Eds.). (1999). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (Forth Edit.). London: Williams and Wilkins.

Noble, W. (2006). Bilateral hearing aids: a review of self-reports of benefit in comparison with unilateral fitting. *International Journal of Audiology*, *45*, S63–71.

Noble, W., Byrne, D., & Ter-Horst, K. (1997). Auditory localization, detection of spatial separateness, and speech hearing in noise by hearing impaired listeners. *The Journal of the Acoustical Society of America*, *102*(4), 2343–52.

Noble, W., & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial, and Qualities of Hearing scale (SSQ). *International Journal of Audiology*, *45*(3), 172–181.

Noble, W., & Perrett, S. (2002). Hearing speech against spatially separate competing speech versus competing noise. *Perception & Psychophysics*, *64*(8), 1325–36.

Nobre, A. C., Sebestyen, G. N., Gitelman, D. R., Mesulam, M. M., Frackowiak, R. S. J., & Frith, C. D. (1997). Functional localization of the system for visuospatial attention using positron emission tomography. *Brain*, *120*, 515–33.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–64.

Nunez, P. L., & Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG* (Vol. 4, p. 611). Oxford University Press.

Ostroff, J. M., McDonald, K. L., Schneider, B. A, & Alain, C. (2003). Aging and the processing of sound duration in human auditory cortex. *Hearing Research*, *181*(1-2), 1–7.

Papso, C. F., & Blood, I. M. (1989). Word recognition skills of children and adults in background noise. *Ear and Hearing*, *10*(4), 235–236.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, *30*(6), 653–61.

Pavani, F., Làdavas, E., & Driver, J. (2003). Auditory and multisensory aspects of visuospatial neglect. *Trends in Cognitive Sciences*, *7*(9), 407–414.

Peterson, G., & Barney, H. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America*, *97*(1), 593–608.

Picton, T. W., & Hillyard, S. A. (1972). Cephalic skin potentials in electroencephalography. *Electroencephalography and Clinical Neurophysiology*, *33*(4), 419–424.

Picton, T. W., & Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Developmental Neuropsychology*, *31*(3), 249–78.

Pittman, A. L., Lewis, D. E., Hoover, B. M., & Stelmachowicz, P. G. (2005). Rapid word-learning in normal-hearing and hearing-impaired children. *Ear and Hearing*, *26*(6), 619–629.

Polich, J. (2009). Updating P300: An Integrative Theory of P3a and P3b, *118*(10), 2128–2148.

Posner, M. I., Inhoff, A. W., Friedrich, F. J., & Cohen, A. (1987). Isolating attentional systems: A cognitive-anatomical analysis. *Psychobiology*, *15*(2), 107–121.

Posner, M. I., Petersen, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of Cognitive Operations in the Human Brain. *Science*, *240*(4859), 1627–1631.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, *109*(2), 160–74.

Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *The Journal of the Acoustical Society of America*, *114*(1), 446.

Rapin, I., Schimmel, H., Tourk, L. M., Krasnegor, N. A., & Pollak, C. (1966). Evoked responses to clicks and tones of varying intensity in waking adults. *Electroencephalography and Clinical Neurophysiology*, *21*(4), 335–344.

Richards, V. M., & Neff, D. L. (2004). Cuing effects for informational masking. *The Journal of the Acoustical Society of America*, *115*(1), 289.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231.

Rose, M. M., & Moore, B. (1997). Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *102*(3), 1768–78.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *336*(1278), 367–73.

Rosen, S., Souza, P. E., Ekelund, C., & Majeed, A. A. (2013). Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *The Journal of the Acoustical Society of America*, *133*(4), 2431–43.

Rudner, M., Lunner, T., Behrens, T., Thorén, E. S., & Rönnberg, J. (2012). Working memory capacity may influence perceived effort during aided speech recognition in noise. *Journal of the American Academy of Audiology*, *23*(8), 577–89.

Rusiniak, M., Lewandowska, M., Wolak, T., Pluta, A., Milner, R., Ganc, M., … Skarżyński, H. (2013). A modified oddball paradigm for investigation of neural correlates of attention: a simultaneous ERP-fMRI study. *Magn Reson Mater Phy*, *26*(6), 511–26.

Sabri, M., Binder, J. R., Desai, R., Medler, D. A., Leitl, M. D., & Liebenthal, E. (2008). Attentional and linguistic interactions in speech perception. *Neuroimage*, *39*(3), 1444–1456.

Salmi, J., Rinne, T., Degerman, A., & Alho, K. (2007). Orienting and maintenance of spatial attention in audition and vision: an event-related brain potential study. *The European Journal of Neuroscience*, *25*(12), 3725–33.

Salmi, J., Rinne, T., Koistinen, S., Salonen, O., & Alho, K. (2009). Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Research*, *1286*, 155–64.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–28.

Salvi, R. J., Lockwood, A. H., Frisina, R. D., Coad, M. L., Wack, D. S., & Frisina, D. R. (2002). PET imaging of the normal human auditory system: responses to speech in quiet and in background noise. *Hearing Research*, *170*, 96–106.

Schadow, J., Lenz, D., Thaerig, S., Busch, N. a, Fründ, I., & Herrmann, C. S. (2007). Stimulus intensity affects early sensory processing: sound intensity modulates auditory evoked gamma-band activity in human EEG. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *65*(2), 152–61.

Schneider, B. a, Daneman, M., & Murphy, D. R. (2005). Speech comprehension difficulties in older adults: cognitive slowing or age-related changes in hearing? *Psychology and Aging*, *20*(2), 261–71.

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, *80*(1-2), 1–46.

Scott, S. K., Rosen, S., Beaman, C. P., Davis, J. P., & Wise, R. J. S. (2009). The neural processing of masked speech: Evidence for different mechanisms in the left and right temporal lobes. *The Journal of the Acoustical Society of America*, *125*(3), 1737–43.

Scott, S. K., Rosen, S., Wickham, L., & Wise, R. J. S. (2004). A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. *The Journal of the Acoustical Society of America*, *115*(2), 813–821.

Shafiro, V., & Gygi, B. (2007). Perceiving the speech of multiple concurrent talkers in a combined divided and selective attention task. *The Journal of the Acoustical Society of America*, *122*(6), EL229–35.

Shamma, S. a, & Micheyl, C. (2010). Behind the scenes of auditory perception. *Current Opinion in Neurobiology*, *20*(3), 361–6.

Shapiro, R., & Janzen, A. H. (1960). *The normal skull: A roentgen study*. New York: Hoeber.

Shinn-Cunningham, B. G. (2005). Influences of spatial cues on grouping and understanding sound. *Proceedings of the Forum Acusticum*.

Shinn-Cunningham, B. G. (2007). Why hearing impairment may degrade selective attention. *Inter. Sym. Auditory and Audiological.*, 1–12.

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182–6.

Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, *12*(4), 283–99.

Shomstein, S., & Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *The Journal of Neuroscience*, *26*(2), 435–9.

Shulman, G. L., d'Avossa, G., Tansy, A. P., & Corbetta, M. (2002). Two attentional processes in the parietal lobe. *Cerebral Cortex*, *12*(11), 1124–31.

Slagter, H. A., Giesbrecht, B., & Kok, A. (2007). fMRI evidence for both generalized and specialized components of attentional control. *Brain Research*, 90–102.

Smith, D. R. R., & Patterson, R. D. (2005). The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *The Journal of the Acoustical Society of America*, *118*(5), 3177.

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Perception of speech in noise: Neural correlates. *Journal of Cognitive Neuroscience*, *23*(9), 2268–79.

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: biological mechanisms. *Cerebral Cortex (New York, N.Y. : 1991)*, *22*(5), 1180–90.

Spence, C. J., & Parise, C. (2010). Prior-entry: a review. *Consciousness and Cognition*, *19*(1), 364–79.

Spierer, L., Meuli, R., & Clarke, S. (2007). Extinction of auditory stimuli in hemineglect: Space versus ear. *Neuropsychologia*, *45*(3), 540–51.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401.

Srinivasan, R., Nunez, P. L., & Silberstein, R. B. (1998). Spatial filtering and neocortical dynamics: estimates of EEG coherence. *IEEE Transactions on Bio-Medical Engineering*, *45*(7), 814–26.

Srinivasan, R., Thorpe, S., Deng, S., Lappas, T., & D'Zmura, M. (2009). Decoding attentional orientation from EEG spectra. In *Human-Computer Interaction. New Trends* (pp. 176–183).

Sussman, E., Winkler, I., & Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin & Review*, *10*(3), 630–7.

Taub, H., & Raab, D. (2005). Fluctuations of N1 Amplitude in Relation to Click-Intensity Discrimination. *The Journal of the Acoustical Society of America, 46,* 969–78.

Tyler, R. S., Summerfield, A. Q., Wood, E. J., & Fernandes, M. A. (1982). Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *72*(3), 740–52.

Varghese, L. A., Ozmeral, E. J., Best, V., & Shinn-Cunningham, B. G. (2012). How visual cues for when to listen aid selective auditory attention. *Journal of the Association for Research in Otolaryngology*, *13*(3), 359–68.

Voisin, J., Bidet-Caulet, A., Bertrand, O., & Fonlupt, P. (2006). Listening in silence activates auditory areas: a functional magnetic resonance imaging study. *The Journal of Neuroscience*, *26*(1), 273–8.

Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *The Journal of Neuroscience*, *23*(13), 5799–804.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The Journal of Neuroscience*, *32*(40), 14010–21.

Wilson, F. A. W., O'Scaoaidhe, S. P. O., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, *260*(5116), 1955–1958.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(18), 8722–6.

Woldorff, M. G., Hansen, J. C., & Hillyard, S. A. (1987). Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. *Current Trends in Event-Related Potential Research*, 146–154.

Woldorff, M. G., Hazlett, C. J., Fichtenholtz, H. M., Weissman, D. H., Dale, A. M., & Song, A. W. (2004). Functional parcellation of attentional control regions of the brain. *Journal of Cognitive Neuroscience*, *16*(1), 149–65.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, *79*(3), 170–91.

Woodrow, H. (1914). The measurement of attention. *The Psychological Monographs, i*.

Woodruff, P. W. R., Benson, R. R., Bandettini, P. A., Kwong, K. K., Howard, R. J., Talavage, T. M., … Rosen, B. R. (1996). Modulation of auditory and visual cortex by selective attention is modality-dependent. *NeuroReport*, *7*, 1909–1913.

Woods, D. L., Alain, C., Covarrubias, D., & Zaidel, O. (1993). Frequency-related differences in the speed of human auditory processing. *Hearing Research*, *66*, 46–52.

Woods, D. L., Hillyard, S. A., & Hansen, J. C. (1984). Event-related brain potentials reveal similar attentional mechanisms during selective listening and shadowing. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(6), 761–777.

Wu, C.-T., Weissman, D. H., Roberts, K. C., & Woldorff, M. G. (2007). The neural circuitry underlying the executive control of auditory spatial attention. *Brain Research*, *1134*(1), 187–98.

Xiang, J., Simon, J. Z., & Elhilali, M. (2010). Competing streams at the cocktail party: Exploring the mechanisms of attention and temporal integration. *The Journal of Neuroscience*, *30*(36), 12084–93.

Yamaguchi, S., Tsuchiya, H., & Kobayashi, S. (1994). Electroencephalographic activity associated with shifts of visuospatial attention. *Brain*, *117*, 553–62.

Yantis, S. (2005). How visual salience wins the battle for awareness. *Nature Neuroscience*, *8*(8), 975–7.

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*, *5*(10), 995–1002.

Yost, W. A., Dye, R. H., & Sheft, S. (1996). A simulated “cocktail party” with up to three sound sources. *Perception & Psychophysics*, *58*(7), 1026–36.

Zampini, M., Brown, T., Shore, D. I., Maravita, A., Röder, B., & Spence, C. J. (2005). Audiotactile temporal order judgments. *Acta Psychologica*, *118*(3), 277–91.

Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., … Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), 2293–8.

Zurek, P., & Delhorne, L. (1987). Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment. *The Journal of the Acoustical Society of America*, 1548–1559.