Neuropsychologia 51 (2013) 2245–2250

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Neural basis of auditory expectation within temporal cortex

J.M. Nazimek [a,*], M.D. Hunter [a], R. Hoskin [a], I. Wilkinson [a,b], P.W. Woodruff [a]

[a] *Sheffield Cognition and Neuroimaging Laboratory (SCANlab), Academic Clinical Psychiatry, Department of Neuroscience, Faculty of Medicine, Dentistry and Health, University of Sheffield, Yorkshire, UK*

[b] *Academic Radiology, Faculty of Medicine, Dentistry and Health, University of Sheffield, Yorkshire, UK*

ARTICLE INFO

Article history:
Received 5 January 2013
Received in revised form
8 July 2013
Accepted 24 July 2013
Available online 6 August 2013

Keywords:
Auditory perception
Predictive coding
Expectation
Prediction error
Associative learning

ABSTRACT

Predictive coding frameworks of perception propose that neural networks form predictions of expected input and generate prediction errors when the external input does not match expectation. We therefore investigated the processing of unexpected sounds and silence in the auditory cortex using fMRI. Unexpected sounds, when compared to expected sounds, evoked greater activation in large areas of the left temporal and insular cortices. Additionally the left middle temporal gyrus exhibited greater activation to unexpected events in general, whether sounds or silence, when compared to the corresponding expected events. These findings support predictive coding models of perception, which suggest that regions of the temporal cortex function to integrate sensory information with predictive signals during auditory perception.

1. Introduction

The predictive coding model of perception (Rao & Ballard, 1999) posits that normal perception consists of an interaction between two types of processing: top-down (expectation driven) and bottom-up (stimulus driven). According to this framework neural networks enable the learning of associations between perceived stimuli and the formation of predictions relating to incoming input. For example, the sound of an engine triggers the prediction of an approaching vehicle. Forming predictive representations of incoming signals is likely to conserve neural resources by reducing the amount of processing required when the expected stimulus occurs (Bubic, von Cramon, & Schubotz, 2010). However, the ability to adapt to a rapidly changing environment relies upon recognition of discrepancies between expected and actual stimulation, and the updating of future expectations accordingly (Friedman, Goldman, Stern, & Brown, 2009). For instance, detection of unexpected sounds in a car engine may help identify a mechanical problem. At a neural level these functions are most likely achieved through a hierarchical structure where each processing level in the brain transmits the expectation to a lower level via feedback pathways.The lower level then compares this expected input to the actual, external stimulation received and returns a residual error signal (the prediction error) communicating any difference between the two, to the higher level via feedforward pathways.

Mismatch negativity (MMN) refers to an enhanced brain wave potential evoked by a deviant sound in a string of standard sounds (Naatanen, Simpson, & Loveless, 1982). The MMN is believed to represent the prediction error signal (Friston, 2005). Functional MRI studies show that the network involved in processing deviant sounds involves superior and middle temporal gyri (Kim et al., 2009; Opitz, Rinne, Mecklinger, von Cramon, & Schroger, 2002; Schall, Johnston, Todd, Ward, & Michie, 2003). However, MMN is usually generated via oddball paradigms, where expectation of a stimulus is generated through repetition (i.e. the stimulus predicts itself). In contrast, 'real world' associations are often a consequence of more complex patterns of stimulation, where one event predicts another, different, event (Bendixen, Prinz, Horvath, Trujillo-Bareto, & Schröger, 2008; Bendixen, Schröger, & Winkler, 2009; Vuust, Ostergaard, Pallessen, Bailey, & Roepstorff, 2009). Furthermore, it has been argued that the MMN generated from oddball paradigms may not reflect a distinct prediction error signal, but instead may be partly or totally a result of stimulus adaption (i.e. reduced neural responding over time) to the repeated standard sounds (May & Tiitinen, 2010). Hence, a paradigm involving learning associations between two different stimuli might represent a more appropriate methodology for studying predictive coding. Such a paradigm could be based on the associative learning theories, which propose that once the association between a cue and the cued outcome has been learnt, unexpected outcomes produce an error signal (Rescorla & Wagner, 1972; Schultz & Dickinson, 2000).

Previous functional MRI studies have demonstrated prediction error as increased activity in brain areas engaged in learning associations, e.g. prefrontal cortex in cognitive tasks and visual cortex in audio-visual tasks (den Ouden, Friston, Daw, McIntosh, &

* Corresponding author. Tel.: +44 121 301 5440.
E-mail address: J.Nazimek@bradford.ac.uk (J.M. Nazimek).

0028-3932/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2013.07.019

Stephan, 2009; Fletcher et al., 2001). Furthermore, this increase in activity is greater for unexpected events (which follow cues that have previously been presented on their own) than for unexpected omissions (which occur when an expected event does not happen; Fletcher et al., 2001). We therefore investigated the process of auditory predictive coding in healthy individuals using fMRI to examine how violations of expectation, including both unexpected events and omissions, are processed in the auditory cortex. We developed an associative learning task consisting of pairs of stimuli and contrasted conditions of expected and unexpected sounds and silences. When cued by a visual stimulus, participants indicated whether an auditory stimulus was sound or silence. Whilst the majority of the auditory stimuli were presented in the context of their previously learnt pairs (i.e. they were expected), the minority appeared in mismatched pairs (i.e. they were unexpected). Behavioural data confirmed that participants were able to learn the association between visual and auditory stimuli when exposed to this design. We hypothesised that once participants learnt associations between the visual and auditory stimuli, auditory events that violate the learnt expectation would evoke greater activity in the auditory cortex than items that match expectation, even in situations where the auditory stimulus is silence. Specifically, we predicted that, following the visual cue, unexpected silence (i.e. silence when the cue had previously been associated with sound) would induce greater activity in auditory cortex than an expected silence. Similarly, an unexpected sound (i.e. a sound when the cue had been previously associated with silence) would evoke greater activity in the auditory cortex than an expected sound. Finally, the increase in activity in the auditory cortex by unexpected sound would be greater than the increase evoked by unexpected silence.

2. Method

2.1. Participants

The study was approved by the School of Medicine Research Ethics Committee at the University of Sheffield. Twelve healthy volunteers (6 men, mean age 24.3 years, SD 5.5 years) with no history of psychiatric or neurological disorders were recruited from the student population of the University of Sheffield. All participants were right-handed and had good hearing, as assessed by The Edinburgh Handedness Inventory (Oldfield, 1971) and Hearing Handicap Inventory for Adults (Newman, Weinstein, Jacobson, & Hug, 1990) Participants gave their informed consent prior to commencing the study.

2.2. Procedure

The associative learning task involved participants being presented with one of two shapes (vertical or horizontal rectangle) and after a 500 ms gap, with both shapes side-by side. They were required to respond as to whether the initially presented shape appeared on the left or the right. Following a further 500 ms pause the words 'sound' and 'silence' appeared side-by-side on the screen accompanied by either the presence or absence of an auditory stimulus (a sinusoidal tone). Participants were required to indicate whether the text describing the actual auditory stimulation appeared on the left or the right. Unbeknownst to the participants, a vertical rectangle was always followed by a tone (30 trials) whereas a horizontal rectangle was always followed by silence (30 trials). Trials were separated by 1 s and their order was randomized (Fig. 1). Responses ensured that participants paid attention to the stimuli, thus increasing the likelihood that they would learn the associations. The associative learning task took approximately 5 min. Initially, to confirm the ability of the paradigm to induce associative learning, a pilot study was conducted using a separate sample to that involved in the scanning study. The results from this pilot study revealed the predicted main effect of expectation, with responses being faster to expected stimuli ($F(1,59)=2.9$, $p<.05$). This confirmed that the paradigm was capable of inducing learning of an association between the visual and auditory stimuli. After completing the task, participants were asked to select which of three sentences correctly described the contingencies between the visual and auditory stimuli. One sentence stated that the horizontal rectangle was most likely to be followed by a sound, the second that the vertical rectangle was most likely to be followed by a sound, and the third that both shapes were equally likely to be followed by a sound. Across the sample the correct statement was only selected at chance levels (35%). This suggests that the participants had remained unaware of the contingencies during the task, and therefore that the learning of the association was implicit.

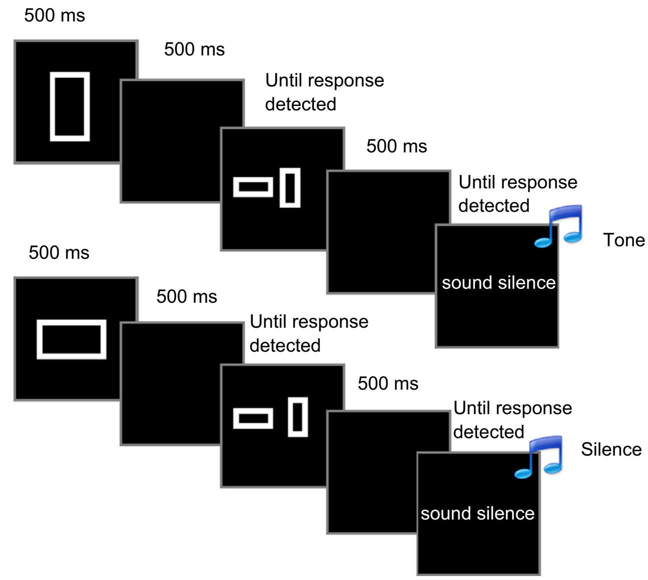

Fig. 1. Task design in behavioural phase. Participants responded to both visual and auditory stimuli. The vertical and horizontal rectangles appeared simultaneously and remained on the screen until a response was detected. Participants were asked to press the right arrow key if the shape on the right corresponded to the previously presented single shape, or the left arrow key if the shape on the left corresponded to previously presented single shape. This part of the task was the same in the scanning phase, in which participants pressed the third button on the response box for the 'right' judgement and the second button on the response box for the 'left' judgement. Response to the sound (required in the behavioural phase only) was cued by words 'sound' and 'silence' appearing in the centre of the screen simultaneously with auditory events until response was recorded. Participants pressed the right arrow key when the word on the right from the centre of the screen corresponded to what they heard and the left arrow key if the word on the left from the centre of the screen corresponded to what they heard. The position of the shapes and words on the screen was randomized.

The MRI study consisted of two phases: a behavioural phase, which was identical to the associative learning paradigm described above, and a scanning phase. The scanning phase began with a refresher session (before the scanning itself, but while the participant was in the scanner). Participants were presented with 20 randomized repetitions of the associations from the behavioural stage: vertical rectangle-tone (10 trials) and a horizontal rectangle-silence (10 trials). This was performed to reinforce the learnt associations from the behavioural phase in response to the time taken to get the participant into the scanner. The refresher was followed by 3 blocks of 16 trials (48 trials altogether), in which expected sounds and silences were contrasted with unexpected sounds and silences. In 25% of trials the visual cues were followed by an unexpected outcome, i.e. the vertical rectangle was followed by silence and the horizontal rectangle was followed by sound. Hence, each block of 16 trials consisted of the following pairings: 6 repetitions of vertical rectangle-tone (expected sound), 6 repetitions of horizontal-rectangle-silence (expected silence), 2 repetitions of vertical rectangle-silence (unexpected silence) and 2 repetitions of horizontal rectangle-sound (unexpected sound). The order of the trials was pseudo-randomized so that the unexpected auditory stimuli did not appear in the first 4 trials. The pause between presentation of the pair of shapes and an auditory stimulus was set at 750 ms. During the scanning phase participants were required to respond to the shapes in the same way as in the behavioural phase, however unlike in the behavioural phase, participants were not required to respond to the sounds. Each block of experimental trials was separated by a block of 8 baseline trials, during which participants were presented with a fixation cross and asked to respond to it by pressing the second button on the response box with their index finger, to control for the motor action in response to the shape in the experimental trials. The fixation cross appeared at the same time point as the pair of shapes in the experimental blocks. In addition to experimental and baseline trials, the first time-point of each block presented participants with instructions as to the task they should perform during that block. This part of experiment took 14 min 45 s.

2.3. Stimuli

Stimuli were presented using commercially available 'Presentation' software (Neurobehavioral Systems, Albany, USA, http://www.neurobs.com). In the behavioural stage, visual stimuli were presented on a computer screen and auditory stimuli were delivered over the headphones. In the scanner visual stimuli were