

Auditory temporal-envelope processing in high-functioning children with Autism Spectrum Disorder

José Ignacio Alcántara*, Thomas E. Cope¹, Wei Cope², Emma J. Weisblatt³

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kingdom

ARTICLE INFO

Article history:

Received 15 April 2011

Received in revised form 27 January 2012

Accepted 31 January 2012

Available online 13 February 2012

Keywords:

Autism Spectrum Disorder

Auditory perception

Temporal-envelope resolution

Processing efficiency

ABSTRACT

Individuals with Autism Spectrum Disorder (ASD) perform worse than controls when listening to speech in a temporally modulated noise (Alcántara, Weisblatt, Moore, & Bolton, 2004; Groen et al., 2009). The current study examined whether this is due to poor auditory temporal-envelope processing. Temporal modulation transfer functions were measured in 6 high-functioning children with ASD and 6 control listeners, using sinusoidal amplitude modulation of a broadband noise. Modulation-depth thresholds at low modulation rates were significantly higher for the ASD group than for the Control group, and generally higher at all modulation rates tested. Low-pass filter model estimates of temporal-envelope resolution and temporal-processing efficiency showed significant differences between the groups for modulation-depth threshold values at low modulation rates. Intensity increment-detection thresholds, measured on a subset of individuals in the ASD and Control groups, were not significantly different. The results are consistent with ASD individuals having reduced processing efficiency of temporal modulations. Possible neural mechanisms that might underlie these findings are discussed.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Autism Spectrum Disorders (ASDs) are characterised by the early onset of a triad of behavioural symptoms: (1) qualitative impairments in social interaction, (2) communication deficits, and (3) restricted, repetitive, and stereotyped patterns of behaviours, activities, and interests (American Psychiatric Association, 1994). Relatively insignificant in the diagnostic criteria are sensory symptoms (*i.e.*, hypo- and hyper-sensitivity to sensory stimuli), which are, however, frequently reported clinically (*e.g.*, Goldfarb, 1963; Iarocci & McDonald, 2006; Kanner, 1943), are distressing (*e.g.*, Grandin & Scariano, 1986), and seem particularly prevalent in the auditory domain (Dahlgren & Gillberg, 1989; Hermelin & O'Connor,

1970; Ornitz, 1974; Rosenhall, Nordin, Sandström, Ahlsen, & Gillberg, 1999).

The relative paucity of information on auditory processing in ASD is surprising given how frequently auditory-related symptoms are reported during the diagnosis process. A recent analysis of Autism Diagnostic Interview Revised (ADI-R) sensory item scores, from a UK database collected by the Oxford-based International Molecular Genetic Study of Autism Consortium (IMGSAC), as part of the diagnostic and phenotypic data collection, showed that in 486 ASD children, 89.3% scored on one or more sensory items, 40.9% had idiosyncratic negative (adverse) sensory reactions, 67.9% had sensory interests, and 65.8% were hyper-sensitive to noise. Moreover, our own further detailed analysis of all available sources (*i.e.*, ADI-R, Autism Diagnostic Observation Schedule, and clinical reports) on 248 UK probands, from the same (IMGSAC) database, revealed that almost all individuals (96.8%) had a sensory feature when all available information was explored. Positive-visual and negative-auditory sensory features were the most common, with cross-modal reactions being rare; negative auditory reactions to background signals occurred in 76.8% of the children. Similarly, a study examining patterns of sensory abnormalities in children and adults with autism using the Diagnostic Interview for Social and Communication Disorders (Leekam, Libby, Wing, Gould, & Taylor, 2002) showed that over 90% of children displayed sensory abnormalities and had sensory symptoms in multiple sensory domains (Leekam, Nieto, Libby, Wing, & Gould, 2007).

* Corresponding author. Tel.: +44 01223 764412; fax: +44 01223 333564.

E-mail addresses: jia10@cam.ac.uk (J.I. Alcántara),

thomas.cope@newcastle.ac.uk (T.E. Cope), wei.cope@nuth.nhs.uk (W. Cope),

ejw44@cam.ac.uk (E.J. Weisblatt).

¹ Present address: Auditory Group, Clinical Laboratory, Institute of Neuroscience, University of Newcastle upon Tyne, Framlington Place, Newcastle upon Tyne NE1 7RU, United Kingdom.

² Present address: Haematological Sciences, Institute of Cellular Medicine, Newcastle University, Framlington Place, Newcastle upon Tyne, NE2 4HH, United Kingdom.

³ Present address: Department of General and Adolescent Paediatrics, Institute of Child Health, University College London, 250 Euston Road, London NW1 2PQ, United Kingdom.

It is likely that reported auditory sensory features are related to the difficulties in understanding speech in the presence of background sounds (e.g., competing speech), which are also frequently reported by individuals with ASD. In order to quantify these perceptual difficulties, Alcántara et al. (2004) assessed speech-in-noise recognition in high-functioning individuals with ASD. The target speech was a sentence taken from the Audiovisual Sentence List material (MacLeod & Summerfield, 1990). Each sentence had a simple syntactic structure, with three keywords for scoring (e.g., “the mother cooked the dinner”). Background noises were filtered to have the same long-term spectra as the target speech, whilst containing fluctuations in amplitude over time (i.e., “temporal dips”) and fluctuations in amplitude across frequency (i.e., “spectral dips”), such as might be found in the background of a typical acoustic environment (e.g., classroom settings), and have previously been associated with improved levels of speech recognition, relative to steady-state backgrounds (Duquesnoy, 1983; Festen & Plomp, 1990).

Some real-life examples might be helpful to the reader at this point to illustrate what is meant by temporal and spectral dips. It is perhaps easier to intuit the role that temporal dips play in real-life communication, by recalling that when we try to communicate in a room full of people speaking, there will be times when the overall level of the background speech is lower than at other times (corresponding to a temporal dip), and the voice of the person we are trying to communicate with stands out and is clearly audible. Spectral dips are somewhat more abstract in nature, and their beneficial effect on speech-in-noise recognition is not so immediately apparent. A common-day example of the effectiveness of spectral dips in improving speech perception can be illustrated from the experience that a radio broadcaster's voice appears clearer and easier to understand in a car when the car radio's treble setting is turned to a high setting. This is the case because the power spectrum of the background sound of a car in motion comprises higher levels of energy in the low frequencies than in the high frequencies (i.e., relatively, there is a ‘spectral dip’ in the high frequencies). Similarly, the long-term average spectrum of speech contains more energy at low frequencies than at higher frequencies. Consequently, by emphasising the high frequencies of the broadcaster's voice, the listener is in effect improving the signal-to-noise ratio, by placing more of the speech energy of the radio broadcaster's voice within the region of the spectral dip of the car noise. It is the case that normal speech and everyday background sounds, including other voices, contain both temporal and spectral dips. Consequently, there will be times when the target speech coincides with a dip in amplitude of the background (i.e., a temporal dip), and we will be able to hear the speech clearly; similarly, at any given time, some frequencies of the background may be high in amplitude, whilst other frequencies of the background are low in amplitude (i.e., a spectral dip); at these times, the frequencies of the speech that will be most clearly audible will be those that coincide with the spectral dips of the background.

The speech-shaped background sounds used by Alcántara et al. (2004) were: (1) a stationary (invariant) noise, (2) a single female talker, containing fluctuations in amplitude over time and across frequency (i.e., temporal and spectral dips), (3) a temporally modulated noise, containing varying speech-like fluctuations in amplitude over time (i.e., temporal dips), (4) a spectrally modulated noise, containing (deterministic) fluctuations in amplitude across frequency (i.e., spectral dips), and (5) a spectro-temporally modulated noise containing temporal dips as in (3) and spectral dips as in (4). The results were expressed in terms of speech reception thresholds (SRT), defined as the signal-to-noise ratio at which 50% of speech is correctly identified. In both the ASD and Control groups, the SRT values for backgrounds that contained temporal modulations were lower (i.e., better) than the SRTs for the unmodulated background. This is consistent with previous

speech-in-noise studies that have found lower SRTs for temporally modulated backgrounds than for unmodulated backgrounds (e.g., Duquesnoy, 1983; Festen & Plomp, 1990; Gustafsson & Arlinger, 1994; Miller & Licklider, 1950; Nelson, Yin, Carney, & Nelson, 2003; Qin & Oxenham, 2003; Takahashi & Bacon, 1992). This phenomenon is well known and is referred to as ‘masking release’, to reflect that although the unmodulated and modulated backgrounds are presented at the same level, the modulated background is not as effective a masker of the speech as the unmodulated background. Although SRTs for the unmodulated background were similar in both groups, significant differences in SRT between the groups were found for backgrounds that contained temporal modulations. Specifically, the ASD group had SRT values for the temporal-modulated background that were 2–4 dB higher than those of the Control group. Thus, the reduced levels of masking release of the ASD group suggested that these participants did not appear to derive as much benefit as the Control-group participants from the fluctuations in amplitude over time present in the background. It is worth noting that the magnitude of the reduction in masking release observed in the ASD group is likely to be sufficient to impact upon everyday-life communication challenges, such as listening to speech in a noisy classroom setting, and might be sufficient to interfere with the development of phonological categories in children with an immature auditory system, and/or neurodevelopmental disorders (Talcott et al., 2002).

Several mechanisms or factors have been proposed to explain the masking-release phenomenon (for a review, see Füllgrabe, Berthommier, & Lorenzi, 2006). One that appears to be particularly important, at least for the identification of speech in noise, is the process of dip listening, whereby the listener takes advantage of the relatively short temporal minima in the fluctuating background to detect portions of the speech signal (e.g., Alcántara & Moore, 1995; Alcántara, Holube, & Moore, 1996; Cooke, 2006; Festen & Plomp, 1990; Füllgrabe et al., 2006; Miller, 1947; Miller & Licklider, 1950). It seems reasonable to assume that those cues used by the auditory system to code temporal-envelope modulations, and the extraction of those cues at higher-level processing stages, will need to be optimal if dip listening is to be a maximally effective listening strategy. Thus the amount of masking release observed in listeners for backgrounds with temporal modulations might be expected to be dependent on two factors; specifically, (1) the accuracy of peripheral temporal cue encoding (i.e., temporal-envelope resolution), if the background fluctuations in amplitude over time are to be accurately tracked in order to extract target speech at times when the amplitude of the background is lowest, and (2) the level of efficiency of a more central auditory system site that extracts cues that code changes in the temporal-envelope and determine the saliency of modulated sounds (e.g., George, Festen, & Houtgast, 2006; Lorenzi, Husson, Adroit, & Debrulle, 2006).

Accordingly, one possible explanation for the difference in masking release previously reported by Alcántara et al. (2004) is that the dip-listening abilities of the ASD group were not as great as those of the Control group as a result of reduced levels of temporal-envelope processing in the ASD group. However, the use of sentence speech material in Alcántara and colleagues' speech-in-noise study impeded an interpretation of the group masking-release difference observed as originating from purely a bottom-up perceptual deficit, as the potential contribution of group differences in language-mediated top-down influences was not controlled for in this study.

A more recent study investigating spectral and temporal “integration”⁴ in a group of high-functioning individuals with

⁴ The authors use the term “integration” in a cognitive manner, to mean the combination of the partial segments of speech information that are audible and perceived to realise the intended word. With respect to auditory processing, auditory

autism and matched-control participants sought to minimise the contribution of possible group differences in top-down linguistic processing abilities by employing word stimuli with equal semantic complexity (Groen et al., 2009). A similar pattern of results as previously reported by Alcántara and colleagues was observed: SRTs measured for speech in unmodulated and spectral-modulated backgrounds were very similar across groups; however, SRTs for the temporal-modulated background condition were significantly poorer in the autism group (Groen et al., 2009). In other words, any differences in top-down linguistic processing that might be present between individuals with ASD and typically developing controls do not appear sufficient to explain the selective deficits in speech recognition in temporal-modulated backgrounds that have been observed in ASD, and appear more consistent with group differences in bottom-up auditory temporal processing.

Two bottom-up temporal-processing factors could in principle account for the speech-in-noise deficits observed previously in ASD: (1) temporal-envelope resolution, and (2) temporal-envelope processing efficiency. However, the results of the two previous speech-in-noise studies reflect the combined effects of temporal-envelope resolution and temporal-envelope processing efficiency. Moreover, the use of a 10-Hz modulation rate in the study of Groen et al. (2009), which contains temporal dips of about 50 ms in duration, was too slow to detect anything but a very serious deficit in temporal-envelope resolution. In contrast, the presence of reduced temporal-envelope processing efficiency in the ASD group could affect performance at any modulation rate, and so may have been the more important determinant of the SRTs observed in the study of Groen et al. (2009).

Although some primary auditory abilities have been investigated in ASD [absolute thresholds (Gravel, Dunn, Lee, & Ellis, 2006); frequency selectivity (Plaisted, Saksida, Alcántara, & Weisblatt, 2003); pitch discrimination (Bonnell et al., 2003); pitch segmentation (Heaton, 2003); frequency-modulation detection (Samson et al., 2011)], to date no study has directly measured auditory temporal-envelope resolution in this population. The temporal resolution of the auditory system refers to its ability to detect changes in the amplitude over time (*i.e.*, changes in the temporal envelope), such as brief silence gaps in a continuous sound or the repeated modulation of a sound's amplitude.

In contrast, the role of auditory temporal processing in disorders such as developmental dyslexia (DD) and specific language impairment (SLI) has been more extensively investigated. For example, the proposed “fast temporal processing deficit” hypothesis (Merzenich et al., 1996) argues that phonemic/phonological impairments in dyslexia are caused by a more general deficiency in the temporal processing of rapidly presented auditory stimuli (Tallal, 1984). Consistent with this hypothesis are findings that children with DD were not as sensitive to the presence of fast temporal modulations of an amplitude-modulated noise as controls (*e.g.*, Lorenzi, Dumont, & Füllgrabe, 2000; Rocheron, Lorenzi, Füllgrabe, & Dumont, 2002). However, Bishop, Carlyon, Deeks, and Bishop (1999) found no evidence that auditory deficits were necessary for, or a sufficient cause of, language impairments. Similarly, a recent review of the evidence for auditory processing deficits in dyslexia concluded that there were as yet no clear and demonstrable relationships between non-speech and speech deficits, and that although auditory deficits

appeared to be more common in SLI, the association was unlikely to be causal (Rosen, 2003).

In order to account for the backward masking effects previously measured in language-impaired (LI) individuals, Hartley and colleagues have argued in favour of an alternative “processing-efficiency” hypothesis (Hartley & Moore, 2002; Hartley, Hill, & Moore, 2003). Processing efficiency, in this context, refers to all the factors, aside from temporal and spectral resolution, that affect the ability to detect acoustic signals in noise; for example, neural events intrinsic to the auditory system used to extract the information of the signal; the degradative effect of internal neural noise; as well as cognitive factors including attention, memory, and learning (Patterson, Nimmo-Smith, Weber, & Milroy, 1982; Werner & Bargones, 1991). Specifically, it was argued that LI individuals have normal auditory temporal resolution, but require a higher signal-to-noise level to detect the speech signal when it is presented in background noise as a result of a more central deficit in processing of temporal speech information (Hill, Hartley, Glasberg, Moore, & Moore, 2004).

It is the case that biologically relevant sounds such as speech, animal vocalisations and music have distinguishing temporal features that are utilised for effective auditory perception. Common temporal features include sound-envelope fluctuations, which may be modelled in the laboratory as amplitude modulation (AM). It is noteworthy, in this context, that significantly higher thresholds for detection of amplitude modulation, together with corresponding physiological measures of amplitude-modulation detection sensitivity, have also been found in dyslexic individuals, consistent with reduced processing efficiency in this group (Menell, McAnally, & Stein, 1999).

Accordingly, the aim of the current study was to investigate whether there is evidence that might support the hypothesis of a temporal-envelope processing deficit in ASD individuals. The objective was to measure temporal-envelope resolution and temporal-envelope processing efficiency in individuals with ASD and typically developing, age- and IQ-matched controls on a psychoacoustic task of AM detection.

A difficulty encountered in measuring the temporal resolution of the auditory system is that when changes are imposed on the time pattern of a sound, for example, a silent gap is inserted in a narrowband noise, there are generally also associated changes in its magnitude spectrum (“spectral splatter”). That is, the insertion of a silent gap in a tone or narrowband noise has the effect of introducing additional frequencies to the test signal that may be audible and used by the listener to detect the test interval. If this is the case, then performance on the task may be based on the detection of spectral rather than temporal cues. In order to avoid this problem, detection threshold for gaps in a broadband noise have also been used to derive measures of auditory temporal resolution, as the introduction of a gap in a broadband noise does not alter its long-term magnitude spectrum (*e.g.*, Penner, 1977). However, there are also limitations and problems associated with this task, with regards to obtaining an informative and ‘pure’ measure of temporal resolution: the use of a broadband noise does not provide a frequency-specific measure of temporal resolution; also, it is not possible to dissociate the effects of processing efficiency from temporal resolution.

In order to address these issues, we employed a method whereby the threshold for detecting sinusoidal changes in amplitude over time (*i.e.*, deterministic temporal dips) of a broadband (*i.e.*, 20-kHz wide) noise was measured as a function of the rapidity of the changes. A broadband noise was used to avoid the introduction of spectral splatter cues. The sensitivity to temporal amplitude modulations in the noise, as a function of the modulation rate, was assessed by adjusting the depth of the modulation of the noise until it could be detected, giving a measure of the modulation depth at

temporal resolution and auditory temporal integration refer to different processes. Temporal integration refers to the auditory system's ability to collect contiguous information over time to enhance the detection or discrimination of stimuli. As such, it is important for explaining auditory absolute thresholds of stimuli less than 500 ms in duration, and loudness perception phenomena. Temporal resolution, in contrast, refers to the minimum time interval within which different acoustic events can be distinguished. In essence, it gives a measure of how much time the ear takes to process auditory information; that is, how ‘fast’ the ear is.

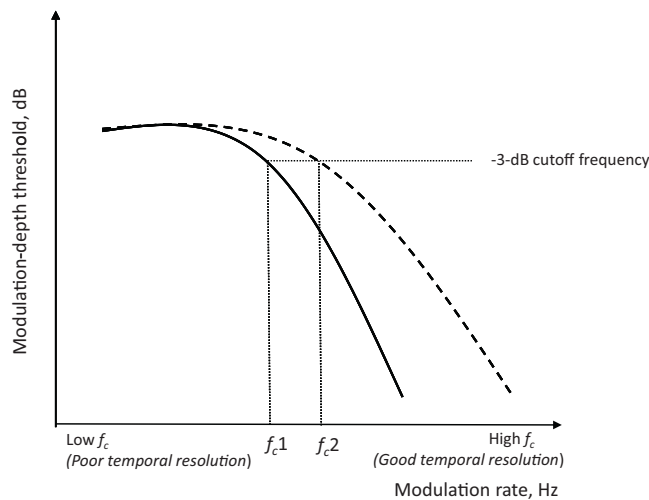


Fig. 1. Schematic representations of differences in temporal resolution in two hypothetical temporal modulation transfer functions. The modulation rate at which modulation-depth thresholds increase by 3 dB, the cutoff frequency (f_c), has a higher value for the dashed-line function than for the solid-line function: temporal resolution is better in the dashed-line function. The modulation-depth thresholds for dashed and solid-line functions are the same at low modulation rates: temporal-processing efficiency is the same for both.

threshold of detection of the amplitude modulation, expressed in decibels (dB).

The function relating the modulation-depth threshold to modulation rate is known as the temporal modulation transfer function (TMTF; Viemeister, 1979). For normal hearing listeners, the modulation-depth threshold is independent of modulation rate for rates up to about 16 Hz. As the rate increases beyond 16 Hz, performance is influenced by the temporal resolution of the auditory system: the value of the modulation-depth threshold increases until at modulation rates above 1000 Hz, modulation cannot be detected at all. In other words, we become progressively less sensitive to the presence of amplitude modulations as the rate of the modulation increases. The overall sensitivity to modulation is reflected in the absolute modulation-depth threshold values, irrespective of modulation rate: low thresholds indicate high sensitivity; high thresholds indicate low sensitivity.

In order to use the TMTF as a performance measure of temporal-envelope resolution, the modulation rate at which the modulation-depth threshold value increases (*i.e.*, gets worse) by 3 dB, called the “cutoff frequency” (f_c), is estimated using the slope of a low-pass filter function used to fit the data. Low f_c values indicate poor temporal resolution; high f_c values indicate good temporal resolution. In addition, the efficiency of temporal processing can be measured by the height of the TMTF, that is, the modulation-depth threshold for slow modulation rates, at which we know temporal resolution does not affect sensitivity.

We predict that if the ASD group has reduced temporal resolution, they will show TMTFs with an f_c value that is lower than that of the Control group, as illustrated in Fig. 1. If, on the other hand, the ASD group is relatively inefficient at extracting features related to temporal modulations, that is, they have reduced temporal-processing efficiency, but normal temporal resolution, then they will display TMTFs with a similar f_c value to the controls, but will have higher modulation-depth thresholds than the Control group, throughout the range of modulation rates tested, as illustrated in Fig. 2.

In order to control for the possibility that individuals with ASD might have reduced stimulus salience, that is, a general auditory-processing efficiency deficit, and consequently perform consistently worse than controls on a range of auditory-perception

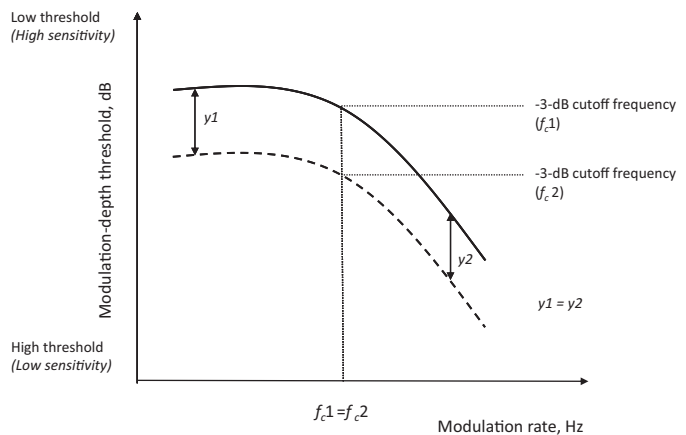


Fig. 2. Schematic representations of differences in temporal-processing efficiency in two hypothetical temporal modulation transfer functions. The modulation rate at which modulation-depth thresholds increase by 3 dB, the cutoff frequency (f_c), is the same value for both the dashed- and solid-line functions: temporal resolution is the same for both. The modulation-depth threshold values for the dashed-line function are higher than for the solid-line function at all modulation rates: temporal-processing efficiency is poorer in the dashed-line function.

tasks, an additional task of intensity discrimination, which was closely matched to modulation detection in procedural terms but did not require the detection of modulation was employed as a follow up study. Three commonly used methods for measuring auditory intensity discrimination are: (1) modulation detection; (2) increment detection; and (3) intensity discrimination of gated stimuli. Although there are some (small) differences in the intensity-difference limens (ΔI) obtained for the different methods, the general trend is similar: for a wideband noise, the smallest detectable intensity change (ΔI) is approximately a constant fraction of the intensity of the stimulus (I). When expressed as the change in decibels, ΔL , then this too is constant, and typically has a value of about 0.5–1 dB for a gated wideband noise presented at levels between 20 and 100 dB SL. Accordingly, an intensity-increment detection task was employed as a follow up study.

2. Materials and methods

2.1. Study 1: Amplitude-modulation detection

2.1.1. Participants

Two groups of participants were used in the study: (i) 6 high-functioning children with confirmed diagnoses of Autism Spectrum Disorder (ASD group; mean age = 12.8 years); (ii) 6 typically developing children (Control group; mean age = 12.5 years).

All ASD participants had received a diagnosis of Asperger's syndrome from specialist autism clinics in the East of England. Diagnoses were confirmed for the purposes of the study by one of the authors (EW), experienced in clinical and research diagnoses of developmental neuropsychiatric disorders. All available sources of information were used, including ADOS and ADI data for participants who had had them performed previously. This included written reports from the clinicians making the diagnosis initially, educational reports and educational and clinical psychological reports, which were supplemented by further history taking from a parent where necessary to make a firm diagnosis for the project. ICD-10 criteria were used, as they are more conservative than DSM-IV criteria. For a diagnosis of Asperger's syndrome, the participants firstly fulfilled criteria for a diagnosis of autism in that they had onset of symptoms before 30 months of age and the required difficulties in the three areas of communication, reciprocal social interaction and repetitive or restricted interests and behaviours. However in addition they were: (1) not late to speak (using ADI criteria of single words by 24 months and phrase speech by 33 months), and (2) had IQ values in the normal range (*i.e.*, >70). The ASD participants were not selected specifically on the basis of auditory symptoms as the interest of this study was in potential auditory processing differences in the ASD population as a whole.

Participants were recruited from the volunteer panel established at the Laboratory for Research into Autism (LARA), and included participants from Cambridgeshire and Bedfordshire in the East of England. They were recruited by direct contact to their parents or guardians if they had previously agreed to participate in the larger series of studies, and through patient support groups in the local area.

Table 1

Individual and mean-group matching data for the Autism Spectrum Disorder (ASD) and typically developing (C) participants. Unpaired, unequal-variance, two-tailed, *t*-test results indicated no significant difference between the groups on measures of Mean age, Verbal IQ, Performance IQ, or Full scale IQ (all $p > 0.10$).

Individual/group	Mean age (years)	Gender	Verbal IQ	Performance IQ	Full scale IQ
ASD1	10.1	Male	121	121	125
ASD2	14.8	Female	99	99	100
ASD3	10.4	Male	151	140	152
ASD4	12.3	Male	108	131	121
ASD5	14.9	Male	131	106	115
ASD6	14.5	Female	121	102	113
ASD mean/SD	12.8/2.0		121.8/16.6	116.5/15.3	121.0/15.9
C1	14.8	Male	114	119	118
C2	14.5	Male	109	109	110
C3	10.9	Male	118	118	120
C4	11.0	Male	104	117	111
C5	11.2	Male	115	129	125
C6	12.9	Female	130	120	128
C mean/SD	12.5/1.6		115.0/8.1	118.7/5.9	118.7/6.6

Groups were matched for mean age, and verbal, performance, and full-scale IQ, as assessed by the Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999), using all four subtests (*i.e.*, vocabulary, block design, similarities, and matrix reasoning). Table 1 summarises the individual and mean participant-matching data. Independent-sample unequal-variance *t* tests (2-tailed) showed no significant differences between the groups on mean age ($p = 0.81$), and verbal ($p = 0.43$), performance ($p = 0.78$), and full-scale ($p = 0.77$) IQ measures. All participants had normal audiometric air-conduction thresholds of 20 dB HL or better at octave frequencies between 250 and 8000 Hz, as measured following the recommendations of the British Society of Audiology (British Society of Audiology, 2004), and were paid for their participation in the study.

2.1.2. Stimuli

A 10-s (20-kHz bandwidth) noise was generated on a Silicon Graphics® UNIX workstation by summing equal-amplitude sinusoids at 1-Hz frequency separation in random phase, and stored on a PC hard drive. For each observation interval, a different 1000-ms segment was selected at random from the generated noise, and was gated on and off with 20-ms raised-cosine ramps. The reference interval comprised an unmodulated noise, $n(t)$; in the target interval, $s(t)$, the temporal envelope of the unmodulated noise carrier was sinusoidally amplitude modulated (SAM) according to Eq. (1):

$$s(t) = c [1 + m \sin(2\pi f_m t + \phi)] n(t) \quad (1)$$

where f_m is the modulation rate (2, 4, 8, 10, 16, 32, 64, 128, 256, or 512 Hz), m is the modulation depth ($0 \leq m \leq 1$), ϕ is the starting phase of the modulation, which was randomised for each interval, and c is a multiplicative compensation term (Viemeister, 1979) set such that the overall power of the noise signals was identical in all intervals. This factor takes into account the increase in average intensity that results from the addition of modulation, which might serve as an additional loudness cue, particularly at higher modulation rates. The expression for c is given in Eq. (2):

$$c = \left[1 + \frac{m^2}{2} \right]^{-0.5} \quad (2)$$

Test signals were digitally generated using a Tucker Davis Technologies (TDT) array processor (TDT AP1) and a 16-bit digital-to-analogue converter (TDT DD1) with a sampling period of 20 μ s (50-kHz sampling rate). Stimuli were low-pass filtered at 20 kHz (Kemo VBF8 Mark 4), passed through a programmable attenuator (TDT PA4) and a headphone buffer (TDT HB6) to a final manual attenuator (Hatfield 2125) set to 10 dB, and presented diotically via Sennheiser HD580 circumaural headphones, which have a diffuse free-field response similar to that of the outer ear. All testing was conducted in a double-walled sound-attenuating booth.

2.1.3. Procedure

Thresholds for the detection of a SAM noise were measured using a 3-interval, 2-alternative forced-choice procedure, and a three-down one-up stepping rule to estimate the modulation depth, m , necessary for 79.4% correct detection (Levitt, 1971). When $m = 1$, the noise is said to be 100% modulated, that is, the depth of the sinusoidal temporal modulations are at a maximum, and therefore easiest to detect; when $m = 0$, it is 0% modulated, and in this case the target interval is effectively the same as the reference interval. The gated noise carrier was unmodulated in two of the intervals, and amplitude modulated in the remaining interval. The silent interval between observation intervals within a trial was 500 ms. Test stimuli were presented at 60 dB SPL. This presentation level was chosen to be clearly audible, yet comfortable to listen to for the duration of the test sessions.

Participants were instructed to indicate the interval containing the “odd-one-out” (*i.e.*, the interval containing the modulated noise) by pressing the corresponding button on a response box, or to guess if they were uncertain of the correct response. Feedback was provided after each trial by a light indicating the correct interval. The step sizes and thresholds were based on the modulation depth in decibels (defined in terms of $20 \log m$). The step size of m variation was 5 dB at the start of each run, and was reduced to 2 dB after the first four reversals. Twelve reversals were obtained in a given run, and the threshold estimate for that run corresponded to the mean of m values of the last eight reversals. At least two estimates were obtained for each modulation-rate condition, with additional estimates collected if the difference between estimates exceeded 3 dB, and/or if the standard deviation of any threshold estimate exceeded 3 dB. The final threshold estimate for a condition was calculated using the average of 2–4 estimates obtained. The order of presentation of modulation rates was randomised across subjects to control for potential order effects. Prior to data collection, each listener received training on SAM noise detection with a 10-Hz modulation rate until modulation-depth thresholds between consecutive runs did not differ by more than 3 dB. In practice, this only required one (2-h) training session for each participant.

All participants attended at least two 2.5-h testing sessions. Within each session, numerous rest breaks were scheduled in order to reduce the possibility of fatigue and loss of interest in the task, including a 20-min break in the middle of the session at which time the WASI test was administered. Participants were informed that they were free to terminate their involvement in the study at any time, should they feel so inclined. The Cambridge Local Research Ethics Committee approved the study, and informed parental consent was obtained before participation of the children in the study.

2.2. Study 2: Intensity-increment detection

2.2.1. Participants

Two groups of participants were used in Study 2, selected from the participants in Study 1 on the sole basis of availability after testing was completed for the first study: 4 high-functioning children with ASD (ASD1, ASD4, ASD5, ASD6; mean age of 11.9 years); and 4 control participants (C2, C4, C5, C6; mean age of 12.5 years). The time interval between Study 1 and Study 2 was about 1 month.

2.2.2. Stimuli

A 10-s (20-kHz bandwidth) noise was generated on a Silicon Graphics® UNIX workstation by summing equal-amplitude sinusoids at 1-Hz frequency separation and in random phase, and stored on a PC hard drive. For each observation interval, a different 500-ms segment was selected at random from the generated noise, and was gated on and off with 10-ms raised cosine ramps. The reference interval comprised a steady-state noise; the test interval comprised the steady-state noise upon which a 200-ms steady-state increment was gated on and off with 2-ms raised cosine ramps, located in the temporal centre of the 500-ms steady state noise.

All other aspects of the generation of the stimuli are as described for Study 1.

2.2.3. Procedure

Increment-detection thresholds were measured using a 3-interval, 2-alternative forced-choice procedure. The gated noise was steady in two of the (reference) intervals, and contained an increment in amplitude in the remaining (test) interval. The silent interval between observation intervals within a trial was 500 ms. The reference interval stimuli were presented at 60 dB SPL. This presentation level was chosen to be clearly audible, yet comfortable to listen to for the duration of the test sessions. The order of presentation of the reference and test intervals was randomised within a run.

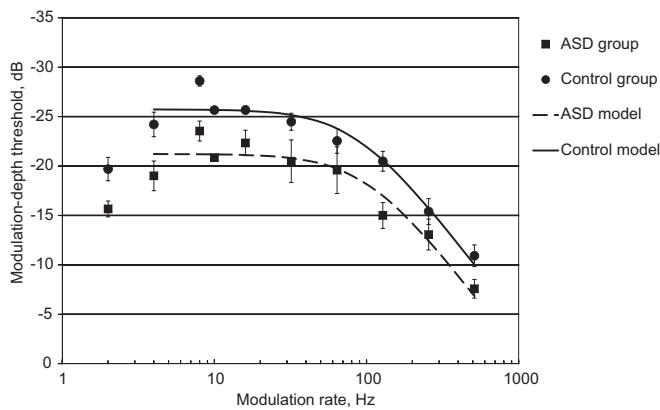


Fig. 3. Mean group modulation-depth thresholds, expressed in decibels (dB), as a function of modulation rate in Hz, are shown for the typically developing (Control) and Autism Spectrum Disorder (ASD) participants (circle and square symbols, respectively), with ± 1 standard error of the mean (SEM). Non-linear least squares regression fits to the modulation-depth data, using a low-pass filter characteristic, are shown for the Control group (solid line) and the ASD group (dashed line).

Participants were instructed to indicate the interval containing the “odd-one-out” (i.e., the interval containing the increment) by pressing the corresponding button on a response box, or to guess if they were uncertain of the correct response. Feedback was provided after each trial by a light indicating the correct interval. The increment was presented at an initial level well above threshold, decreased after three consecutive correct responses, and increased after an incorrect response, to estimate the 79.4% correct point on the psychometric function (Levitt, 1971). Initial step sizes were 2 dB until the fourth reversal, and then 1 dB. The estimated threshold value was the average level of the last 8 reversal points. Nine experimental blocks were run in one testing session lasting 2 h for all test participants, with the exception of ASD4, who only completed the first 8 blocks. The individual-mean increment-detection threshold was the average of the threshold estimates obtained across experimental blocks.

3. Results

3.1. Study 1: Amplitude-modulation detection

Fig. 3 shows the mean modulation-depth thresholds, expressed in decibels ($20 \log m$) for the ASD and Control groups, indicated by the square and circle symbols, respectively. Error bars denote ± 1 standard error of the mean (SEM). The worst (i.e., highest or least negative value) threshold that can be measured is 0 dB, which corresponds to a modulation depth of 1 (100% modulation); more negative values of $20 \log m$ indicate increased sensitivity of temporal modulations.

For each group, the TMTF resembled a typical low-pass filter characteristic, similar in form to those reported by previous authors for typically developing children with normal hearing (e.g., Hall & Grose, 1994; Lorenzi et al., 2000) and adults (e.g., Eddins, 1993). For both the ASD and Control group, modulation-depth thresholds are best for modulation rates of up to about 16 Hz; for modulation rates between about 64 and 128 Hz, modulation-depth thresholds increase by about 3 dB (i.e., become less negative, signifying a reduction in sensitivity to the presence of temporal modulations in level) in both the ASD and Control groups; beyond 128 Hz, the modulation-depth thresholds for both groups increase at a rate of about 3–6 dB/octave. It is also apparent that modulation detection is about 5–6 dB poorer at modulation rates of 2 Hz than at 16 Hz, giving the TMTFs a slight bandpass appearance.⁵

⁵ As previously described in Viemeister (1979), this downward shift for low modulation rates is observed for TMTFs measured with relatively short-duration stimuli and gated noise carriers such as employed in current study. This pattern has been shown to largely disappear with long-duration stimuli (i.e., >1000 ms) or continuous noise carriers, such that modulation-depth thresholds at 2 Hz are similar in level

Of more specific interest is the observation that modulation-depth thresholds for the ASD group were consistently worse than for the Control group, across all modulation rates tested, indicating that the ASD group appeared to be less sensitive than the Control group to the presence of amplitude modulations, regardless of the modulation rate. Specifically, the mean modulation-depth threshold for the ASD group across all modulation rates was -18.2 dB; for the Control group, it was -21.8 dB – a difference of almost 4 dB. Expressed another way, recall that the highest (i.e., poorest) modulation-depth threshold that can be measured is 0 dB, which corresponds to 100% modulation. Accordingly, the modulation-depth thresholds for the ASD and Control groups corresponded to about 12% and 8% modulation, respectively.

A two-way (mixed design) repeated-measures analysis of variance (ANOVA) on the modulation-depth threshold data, with Group (ASD, Control) as a between-groups factor, and Modulation rate (4–512 Hz) as a within-subjects factor showed the following: Group ($F_{1,10} = 4.93$, $p = 0.05$); Modulation rate ($F_{8,80} = 55.0$, $p < 0.001$); Group $\times$ Modulation rate ($F_{8,80} = 0.64$, $p = 0.74$).

Since the form of the TMTF resembles that of the transfer function of a simple low-pass filter (Viemeister, 1979), the modulation detection data for modulation rates of 4–512 Hz⁶ were fitted using a non-linear least-squares regression procedure, based on a low-pass filter characteristic of the form given by the following equation (Eddins, 1993):

$$y = 10 \log_{10} \left\{ \frac{1}{1 + (\alpha f)^2} \right\} + c \quad (3)$$

where c is the y -intercept, α is a constant (the inverse value of which gives the -3 -dB cutoff frequency, f_c), and f is the modulation rate in Hz. The -3 -dB cutoff-frequency value indicates the modulation rate at which there is a reduction in the filter output of 50%, and is a measure of temporal-envelope resolution, expressed in frequency terms (i.e., Hz). Using this fitted function, an associated time constant (τ) was also derived from the modulation-depth threshold data, using Eq. (4), to obtain an operational-equivalent estimate of temporal resolution, expressed in the temporal domain (i.e., seconds):

$$\tau = \frac{1}{2\pi f_c} \quad (4)$$

The low-pass filter model fits to the data are shown in Fig. 3: the solid line indicates the Control-group model fit; the dashed line indicates the ASD-group model fit. The low-pass filter model fitted the data well, and has the added advantage of corresponding to a well-known physical process. The model fits were a significantly better description of the data than the null hypothesis that predicts that modulation-depth threshold data were in fact estimates of the same (global) mean value (ASD group: $F_{1,7} = 78.9$, $p < 0.001$; Control group: $F_{1,7} = 121.1$, $p < 0.001$); and any variations in the data after fitting were small (ASD group: $F_{7,45} = 0.92$, $p = 0.50$; Control group: $F_{7,45} = 2.0$, $p = 0.08$). Thus, there is a high degree of confidence in using this model to estimate the values of the -3 -dB down point (f_c) and the corresponding value of τ , to give an estimation of temporal-envelope resolution, as well as the y -intercept, c , to give an estimation of processing efficiency of temporal modulations.

to those measured at 4- and 16-Hz modulation rates. It is likely that for relatively short-duration stimuli, such as used in the current study, good performance on the task at very low modulation rates becomes more dependent on the intensity discrimination (i.e., amplitude resolution) abilities of the participants rather than by tasks related to using the temporal cues present in the stimuli (Moore, 2003).

⁶ Data for the 2-Hz modulation rate were excluded from the low-pass filter model fits as the use of the 1000-ms test signal in the current study meant that the modulation-depth threshold data for this modulation rate was unlikely to be a pure measure of temporal resolution, for the reasons described in footnote 5.

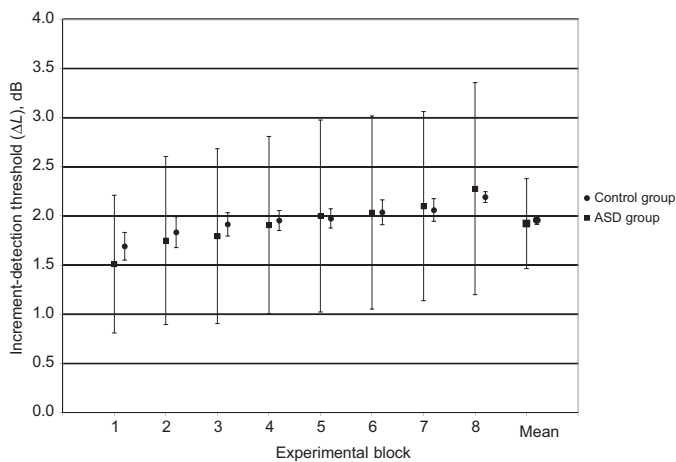


Fig. 4. Mean group increment-detection thresholds, expressed in ΔL (dB), across 8 experimental blocks, and collapsed across experimental blocks (1–8). Square symbols denote the ASD group; circle symbols denote the Control group.

Accordingly, the estimated f_c values were 99.7 Hz for the ASD group, and 85 Hz for the Control group, corresponding to τ values of 1.6 ms and 1.9 ms, respectively. This suggests that the model-estimated temporal-envelope resolution of the ASD group was slightly better than that of the Control group. The model-estimated y-intercept value of the Control group was -25.7 dB, and -21.2 dB for the ASD group, suggesting that the ASD group were less sensitive than the Control group to slow amplitude modulations. Independent-sample unequal-variance t tests (2-tailed) indicated that the estimated y-intercept values were significantly higher (i.e., worse) for the ASD group than they were for the Control group [$t_{10} = 2.47$, $p = 0.03$], and that there was no significant difference between the τ values of both groups [$t_{10} = 1.4$, $p = 0.19$].

In summary, the pattern of elevated modulation-depth thresholds, significantly higher y-intercept values, and similar group τ values (i.e., similar levels of temporal-envelope resolution), suggests that processing efficiency of temporal modulations was reduced in the ASD group, relative to the Control group.

3.2. Study 2: Intensity-increment detection

Inspection of the individual increment-detection thresholds across experimental block indicated that participants ASD1 and C4 had threshold estimates for the final (9th) block that were 1.6 and 1.9 dB, respectively, higher than their more-stable threshold estimates for experimental blocks 1–8. In addition, participant ASD4 did not complete the final experimental block. Consequently, only threshold estimates obtained for the first 8 experimental blocks, where there was a complete set of data for all participants, and relatively stable threshold estimates across experimental blocks, were used for further analyses.

Mean-group increment-detection thresholds as a function of experimental block number (“1–8”), and collapsed across 8 experimental blocks (“Mean”), are shown in Fig. 4. The data points for the Control group have been offset for ease of viewing. Thresholds are expressed in decibels, ΔL , according to Eq. (5):

$$\Delta L = 10 \log_{10} \left\{ I + \frac{\Delta I}{I} \right\} \quad (5)$$

where I is the intensity of the pedestal, and ΔI is the intensity of the increment. According to Eq. (5), the value of ΔL represents the just-noticeable difference in intensity, expressed as the ratio of the higher-level sound to the intensity of the lower level sound, when the listener can just detect a difference between them. Accordingly,

ΔL must always have a positive value, and small ΔL values indicate increased sensitivity to intensity (i.e., loudness) changes.

It is apparent from Fig. 4 that ΔL values for the ASD and Control groups were very similar across the 8 experimental blocks, although there was trend for the size of ΔL to increase with increasing experimental block number, for both participant groups: on average, ΔL values increased from about 1.6 dB to 2.2 dB, from the first to the last experimental block. The pair of square (ASD group) and circle (Control group) symbols on the extreme right of Fig. 4 indicate the mean ΔL values, collapsed across experimental block, and associated standard errors of the means (SEM) for the ASD and Control groups, where $\Delta L = 1.92$ dB (SEM = 0.46) and 1.96 dB (SEM = 0.04), respectively. An independent-sample unequal-variance t test (two-tailed) on the individual mean ΔL values, collapsed across experimental block, indicated there was no significant difference of group ΔL values ($t_6 = 0.08$; $p = 0.94$).

4. Discussion

4.1. Is there evidence for an auditory temporal-processing deficit in ASD?

The main aim of this study was to investigate a temporal processing deficit hypothesis that would be consistent with the reduced masking-release levels for temporally-modulated backgrounds previously observed for individuals with ASD (Alcántara et al., 2004; Groen et al., 2009). A psychoacoustic task of amplitude-modulation detection was employed – the temporal modulation transfer function (TMTF) – that allowed a simultaneous assessment of temporal-envelope resolution and temporal-processing efficiency, both of which might be expected to influence masking release as mediated by a dip-listening strategy. Here, the TMTF describes the modulation depth at threshold for a broadband sinusoidal amplitude-modulated (SAM) noise as a function of its modulation rate. When fitted with a low-pass filter function, the TMTF provides both a measure of temporal-envelope resolution, the -3 -dB down cut-off frequency (f_c), which may also be expressed as a time constant (τ), and a measure of temporal-processing efficiency: the lowest modulation-depth threshold value of the fitted data.

The results showed that the Control group performed in a similar manner to that previously reported for similar-aged children and adults. Specifically, the shape of the TMTF for this group resembled a low-pass characteristic, and the variation in modulation-depth thresholds as a function of modulation rate was well fitted using a low-pass filter function. Although the TMTFs of the ASD individuals similarly resembled a low-pass shape, and were well fitted by the same low-pass filter function, the mean ASD-group modulation-depth threshold values at all modulation rates were higher than that of the Control group, indicating reduced temporal-envelope processing efficiency in the ASD group. The empirical data obtained for the ASD and Control showed a significant (between-group) difference for modulation-depth threshold ($p = 0.05$), and a clear trend of raised modulation-depth thresholds for the ASD group across modulation rates.

It is theoretically possible that the higher modulation-depth thresholds in the ASD group may have been due to reduced levels of top-down sustained and focused attention to the TMTF task, which consistently affected every experimental trial they completed. This would be expected to result in reduced salience of auditory stimuli in general. If this were the case, then we would expect to see generally poorer performance in the ASD group on another auditory task that does not involve detection of temporal modulations, but is closely matched in procedural terms; for example, the intensity-increment detection task employed in Study 2. As

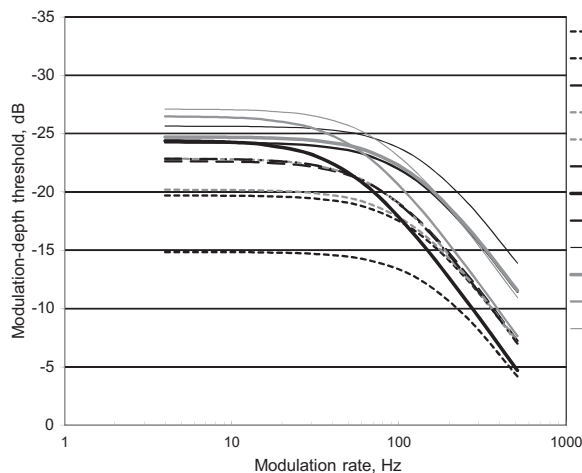


Fig. 5. Low-pass filter model fits for the modulation-detection task (Study 1) for individual participants; solid lines denote Control group; dashed lines denote ASD group.

group-mean increment-detection thresholds were not significantly different ($p = 0.94$), our ASD group appeared to have relatively more difficulty in extracting those cues that specifically code for modulations in amplitude over time (i.e., coding of the temporal envelope), though no observably greater difficulty than the Control-group with the extraction of cues that code for overall level.

Study limitations: Because of the small sample size, caution must be employed in generalising the results of the current study to the wider ASD population until it has been replicated. It is also clear from Figs. 3 and 4 that the inter-subject performance levels for the ASD group were more variable than for the Control group. Increased heterogeneity of performance on a range of behavioural tasks is a well-documented feature for ASD groups. However, given the small sample sizes used in the current study, it is important to examine not only group-mean differences, but also any individual patterns of performance across the two auditory tasks that might suggest an alternative explanation for our results.

Fig. 5 shows the individual-participant low-pass filter model fits of the TMTF data in Study 1. It is apparent that although the ASD individuals exhibited modulation-depth thresholds that were higher than the Control-group participants, some ASD individuals had mean modulation-depth thresholds that approached the normal (Control-group) range, whilst others were more significantly elevated. Given the marginal significance value of the Group factor, and the presence of one individual in the ASD group (ASD2), whose modulation-depth thresholds for low-modulation rates (≈ -15 dB) were about 3-dB higher than the ASD-group mean (-18.2 dB), two additional tests of Group differences in modulation-depth thresholds were carried out to ensure that the significant Group effect found in the ANOVA test was not unduly influenced by the effect of participant ASD2's modulation-detection thresholds on the ASD-group sample mean value. On the basis of the non-significant Group $\times$ Modulation-rate term, modulation-depth thresholds collapsed across modulation rate were contrasted across the ASD and Control groups using a parametric and a non-parametric test; the latter being insensitive to the presence of possible outliers. The results were as follows: Levene's test for equality of variances, $F = 0.982$, $p = 0.345$; independent-sample t -test with equal-variances assumed, $t_{10} = 2.335$, $p = 0.042$ (two-tailed); independent-sample t -test with equal-variances not assumed: $t_{7.39} = 2.335$, $p = 0.05$ (two-tailed); Mann-Whitney U -test, critical $U = 6$; $U = 6$, $p = 0.055$ (two-tailed). Although the Group effect using the non-parametric test was just above the critical significance level, we interpret the very similar p values for the parametric

Table 2

Individual and mean-group increment-detection thresholds expressed in ΔL (dB), collapsed across 8 experimental blocks, and gradient values of regression lines for ΔL as a function of experimental blocks (1–8), for the ASD and Control (C) groups. SD denotes the standard deviation; SEM denotes the standard error of the mean.

Individual/group	ΔL (dB)	SD	Regression line gradient
ASD1	3.15	0.43	0.17
ASD4	0.96	0.16	0.06
ASD5	1.91	0.14	0.06
ASD6	1.66	0.24	0.09
ASD group	1.92	0.46 (SEM)	0.09 (SD = 0.05)
C2	1.93	0.16	0.06
C4	1.87	0.14	0.05
C5	1.95	0.24	0.10
C6	2.08	0.10	0.03
Control group	1.96	0.04 (SEM)	0.06 (SD = 0.03)

and non-parametric tests as an indication that participant ASD2's modulation-detection thresholds did not have an undue effect on the ASD sample mean in the ANOVA results.

Fig. 4 also indicates that there is increased inter-subject variability in increment-detection thresholds for the ASD group, relative to the Control group, as shown by the size of the error bars displayed: standard deviations (SD) of the mean ΔL per experimental block, and standard error of the sample mean ΔL across experimental block. Table 2 shows the individual-mean ΔL values, collapsed across experimental block, and their respective SD values. The source of the increased variability on this task in the ASD group is apparent: participants ASD5 and ASD6 performed at a similar level to the Control group; however, participant ASD1 performed much worse and ASD4 performed much better than the Control group. As the individual-mean thresholds for this task were based on 8 separate threshold estimates, collected over a 2-h period in 8 experimental blocks, it is possible that the larger variability in performance of the ASD group was the result of increased tiredness and/or attentional lapses in the ASD children, that would be exacerbated with increasing repetitions of the task. Accordingly, regression-model lines-of-best fit were calculated for each of the test participants who took part in Study 2, and are shown in Fig. 6: solid lines are for the Control-group children; dashed lines are for the ASD-group children. All of the regression models accounted for a significant proportion of the variability in threshold estimate as a function of experimental block (all R^2 values ≥ 0.7 ; all p values < 0.05). Table 2 (third column) also shows the gradient

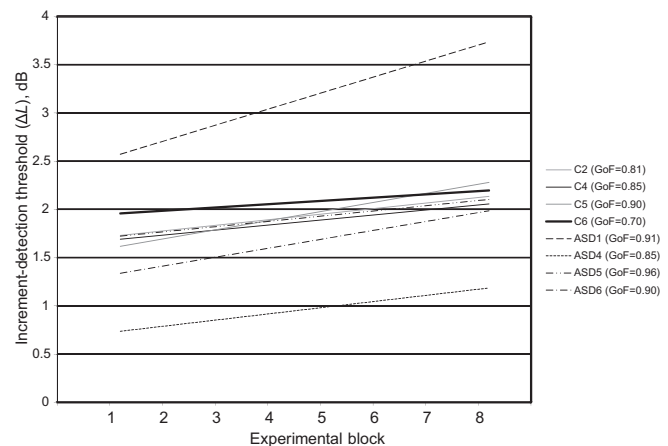


Fig. 6. Linear regression lines for the increment-detection task (Study 2) for individual participants; solid lines denote Control group; dashed lines denote ASD group. "GoF" denotes Goodness of Fit, as estimated by R^2 value.

values of the individual regression lines. Fig. 6 indicates a consistent pattern of a small increase in the increment-detection threshold value with increasing repetitions of the task for all individuals in both the ASD and Control groups. This degradation in performance level with experimental block number was likely to be due to the participants becoming less attentive given the repetitive and uninteresting nature of the task (*i.e.*, showing reduced vigilance). The individual regression-line gradients showed no significant difference between the groups ($t_6 = 1.19$, $p = 0.28$), suggesting that the ASD children were not becoming any less attentive than the Control-group children, during the 2-h period of data collection on this task.

A further possible explanation for the greater variability in individual performance of the ASD children, on both the increment-detection and modulation-detection tasks, is varying levels of processing efficiency for sounds in general in the ASD group. If this were true, then an individual's level of performance on the modulation-detection task should be related to their level of performance on the increment-detection task, as the most likely perceptual effect of reduced processing efficiency would be to reduce the overall salience of the auditory signal. There appears to be some evidence of this in the individual thresholds obtained for the two auditory tasks: participants ASD5 and ASD6 performed relatively well on both tasks; participant ASD1 performed more poorly on both tasks; however, participant ASD4, who had a very similar level of performance on the modulation-detection task to ASD1, also had the lowest (best) increment-detection threshold of the entire cohort, which would be inconsistent with an explanation based on a general processing-efficiency deficit.

It is also possible that the greater spread of individual IQ scores in the ASD group might have had an influence on the level of performance on both auditory tests, if we assume that the cognitive load of both tests is not significantly different. However, there does not appear to be a consistent relationship between IQ score and performance level on the two auditory tasks within individuals. For example, participants ASD1 and ASD4, who both exhibited very similar performance levels for modulation detection, and who had very similar Full-scale IQ scores, performed very differently on the increment-detection task.

Although there does not appear to be a consistent pattern of level of individual performance on the auditory tasks and individual IQ, it would be imprudent to completely rule out the potential influence of individual IQ differences on performance of some auditory-perception tests. A positive relation between IQ and measures of frequency and temporal resolution has previously been reported in children and adults (Hartley, Wright, Hogan, & Moore, 2000). Thus, it is possible that the different IQ scores of the ASD individuals might reflect different levels of sustained attention (Schweizer & Moosbrugger, 2004), which have previously been shown to influence performance on some auditory-perception tasks (Sutcliffe, Bishop, Houghton, & Taylor, 2006). It may therefore be noteworthy that participant ASD2 who performed very poorly on detection of modulation had the lowest Verbal, Performance, and Full-scale IQ scores of the entire cohort. Moreover, participant ASD3, who was amongst the strongest performers within the ASD group, had the highest Verbal, Performance, and Full-scale IQ scores of the entire cohort. Furthermore, in a recent study that investigated the relation between auditory discrimination abilities and auditory sensory behaviours in a large group of ASD individuals ($n = 72$), a sub-group of ASD individuals ($n = 14$) who had better frequency-discrimination thresholds than normal were also more likely to have higher IQ scores (Jones et al., 2009). We shall return to the possible effects of age and IQ on the level of performance on auditory tasks later in this section.

In summary, there does not appear to be a consistent relation between the individual levels of performance observed in

Studies 1 and 2 that strongly suggests that our ASD children suffered from overall reduced salience of auditory signals. Accordingly, we propose an alternative explanation for the reduced modulation-detection thresholds in the ASD group; that is to say, that our ASD children are relatively less efficient at extracting those cues that code for the temporal envelope of sounds. An explanation of reduced processing efficiency for temporal modulations in ASD is also more consistent with previous findings of normal speech recognition in unmodulated and spectrally modulated backgrounds in ASD individuals (Alcántara et al., 2004; Groen et al., 2009). Moreover, individuals with ASD appear to have intensity-discrimination thresholds within the normal range (Jones et al., 2009), and normal levels of processing efficiency for detection of a tone in a simultaneously presented notched-noise (Plaisted et al., 2003), which is inconsistent with reduced salience of sounds in general.

It should, however, be noted that the extent to which any consistent individual patterns of performance across the two studies can be identified is limited by the small sample size of the current study. Accordingly, the remainder of the discussion will examine the extent to which the hypothesis of a temporal processing-efficiency deficit in ASD might be supported from the results of speech-in-noise studies in ASD; how such a deficit might be conceptualised within the framework of a model of auditory temporal processing and related to the development of auditory temporal processing abilities in children; whether the presence of this auditory deficit can be accommodated within existing frameworks of sensory processing in ASD; and whether this auditory deficit is a distinctive feature of ASD, or might also provide an account of the speech-in-noise perception deficits and language problems that have been observed in other language-based learning-disorders, as well as in individuals diagnosed with central auditory processing disorder (CAPD).

4.2. Temporal-envelope processing efficiency and speech-in-noise perception deficits in ASD

The presence of a temporal-processing efficiency deficit in ASD provides a potentially attractive and parsimonious explanation for the observed differences in masking-release level between the ASD and Control groups for backgrounds that contain temporal modulations (Alcántara et al., 2004; Groen et al., 2009). To a first approximation, reduced salience of temporal modulations could be thought of as equivalent to a reduction in the modulation depth. As SRTs have been reported to improve with increases in the magnitude of background fluctuations (Howard-Jones & Rosen, 1993), and similarly, the amount of masking release observed for speech in temporally modulated backgrounds is monotonically related to the modulation depth of a masker (Gnansia, Jourdes, & Lorenzi, 2008), reduced salience of the temporal dips in a background could result in a reduction of the efficacy of employing a dip-listening strategy in a speech-in-noise recognition task, leading to reduced masking release levels for speech in backgrounds with temporal modulations.

It is clear that the strength of an argument that links a temporal processing-efficiency deficit to reduced levels of speech masking release for backgrounds with temporal modulations relies heavily on a dip-listening account of the masking release phenomenon. However, several alternative mechanisms that might also account for masking release have been previously proposed (Füllgrabe et al., 2006), not all of which rely so heavily on good temporal-envelope processing abilities. For example, the observed difference in speech masking release between the ASD and Control-groups might also be the result of differences in the operation of high-level processes involved in, for example, auditory grouping (*e.g.*, Bregman, 1990), perceptual restoration (*e.g.*, Warren, 1970), and/or

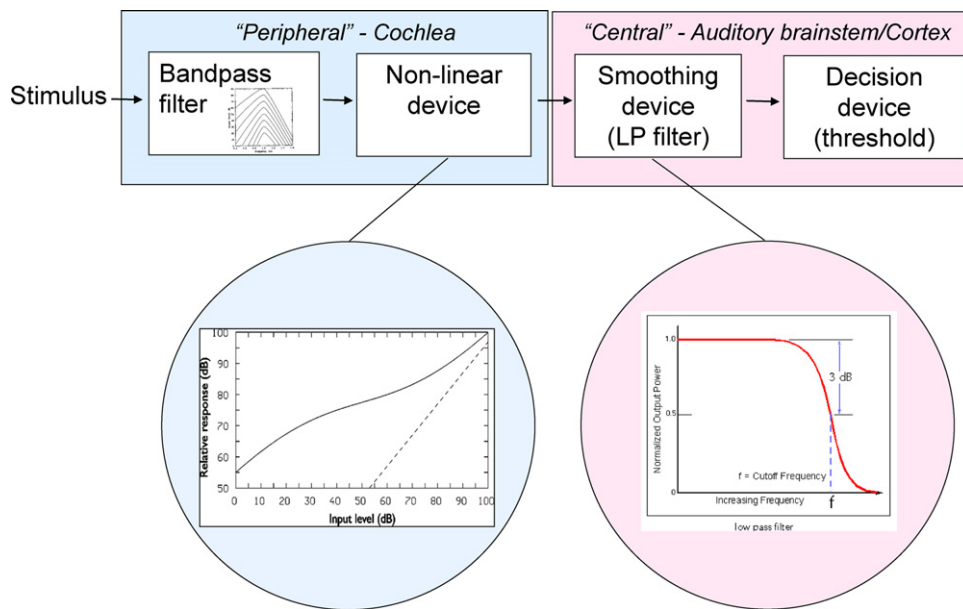


Fig. 7. A block diagram of a model of auditory temporal processing. See text and footnote 4 for further detail.

informational masking (e.g., Summers & Molis, 2004). Moreover, a temporal processing-efficiency deficit would also be expected to reduce the salience of temporal modulations contained within the speech signal itself, which have previously been shown to be important for speech intelligibility (e.g., Drullman, Festen, & Plomp, 1994), and might therefore account for at least a part of the speech-in-noise deficit observed in ASD (though not easily for the differential effects of temporally modulated backgrounds). It is also possible that the temporal modulations in the background might interfere (perceptually) with auditory processing of temporal modulations in the speech envelope (Kwon & Turner, 2001), resulting in a reduction of the efficacy of employing a dip-listening strategy in a speech-in-noise task, when the background contains temporal modulations.

Thus, although the temporal processing-efficiency deficit observed in the current study can be plausibly related to the speech-in-noise deficits observed in ASD, there are also limits to which it can be taken as a comprehensive account of the reduced masking-release levels observed in this group. In the following section, we examine the extent to which the processing-efficiency deficit observed in our ASD children may be related to deficits of relatively low-level auditory processing stages.

4.3. Possible mechanisms underlying poor auditory temporal-envelope processing efficiency in ASD

Most models of auditory temporal-envelope resolution are based on the idea that there is a process at levels higher than the auditory nerve that is 'sluggish' in some way, thereby limiting temporal resolution. Fig. 7 shows a schematic representation of one such model that is similar to those that have been proposed to account for various phenomena involving temporal resolution, such as modulation detection (Viemeister, 1979), gap detection (Forrest & Green, 1987), and some temporal aspects of non-simultaneous masking (Moore, Glasberg, Plack, & Biswas, 1988).⁷

⁷ The auditory temporal-processing model involves four stages: an initial bank of bandpass filters, a nonlinear device, a temporal integrator or smoothing device, and a decision device. Briefly, the initial stage represents the cochlear filtering of sound that occurs in the inner ear. This is the basis for the observed frequency selectivity of the ear, that is, the ability to 'hear out' sounds with different frequencies when they are presented simultaneously (Moore, 2003), and the ability to detect an auditory

signal when it is presented together with background noise of varying bandwidth (Fletcher, 1940). The output of this cochlear filtering is followed by a nonlinear device, which reflects the operation of several processes that occur in the peripheral auditory system, such as neural transduction, whose effects resemble half-wave rectification, that is, only the positive phase component of the stimulating waveform results in neural firing (Young, 2008); and the nonlinear amplitude compression – "compressive nonlinearity" – that occurs on the basilar membrane, whereby low-amplitude signals are amplified more than high-amplitude signals (Sellick, Patuzzi, & Johnstone, 1982). The output of the nonlinear device is fed to a 'smoothing' device, which may be implemented as a low-pass filter (Viemeister, 1979). This has the effect of removing the fast temporal modulations and preserving the slow temporal modulation of output of the nonlinear device such that the output of the low-pass filter resembles the temporal envelope of the output of the nonlinear device. The output of the low-pass filter is fed to the final stage, the decision device. The decision device may use different 'rules' depending on the task required. In the context of the current study, where the task is to detect amplitude modulation of a sound, the device might assess the amount of variance at the output of the low-pass filter (Viemeister, 1979). Alternatively, the ratio of the maximum to the minimum sample at the output of the low-pass filter might be calculated (Forrest & Green, 1987). It is also possible that the threshold for detecting amplitude modulation of a given rate depends on the magnitude of the output of the low-pass filter; the decision device in this case would be equivalent to a simple threshold device (Rodenburg, 1977).

Although all stages of the model can affect temporal resolution, there is a general consensus that the third stage is most directly related to the observed limits of sensitivity of temporal-envelope modulation rate. Processes limiting temporal resolution appear to reflect the operations of auditory brainstem nuclei, and in particular the inferior colliculus (IC), where cutoff frequencies of many neurones approximate the cutoff frequencies typically observed for TMTFs (e.g., Langner & Schreiner, 1988; Schreiner & Langner, 1984). For modulation processing, the operation of the third stage has been characterised as a low-pass filter in the envelope domain (e.g., Viemeister, 1979), with low modulation frequencies passed and high modulation frequencies attenuated, allowing estimation of temporal envelope resolution, f_c . As the f_c estimates for the ASD and Control groups were not found to be significantly different, it would appear that the third stage of the temporal-processing model is not atypical for the ASD individuals.

The fourth stage of the temporal-processing model – the decision device – reflects the operation of extraction of those features that relate to temporal modulation and determine overall sensitivity to modulation. The IC may also be involved in this process, as it appears to play a transitional role between the temporal

representation of AM in the periphery (Joris & Yin, 1992, 1998) and a more rate-based code in the cortex (Liang, Lu, & Wang, 2002). Specifically, the observed temporal resolution of IC neurones is decreased by an order of magnitude or more, relative to the auditory nerve, and neural encoding of AM most likely shifts from a temporal code to a more-robust rate code, which might be more amenable to the operation of the decision device (Frisina, 2001).

In computational versions of the temporal-processing model, the decision device has been instantiated as a simple threshold device that is based on the magnitude of the output of the low-pass filter stage (Rodenburg, 1977) or the ratio of the variance of the modulated and unmodulated sounds (Viemeister, 1979). In accounting for thresholds in forward, backward, and combined forward- and backward-masking conditions, Oxenham and Moore (1994) implemented a decision device that was based on the largest difference in output (of a sliding temporal integrator) between the output of the interval containing the masker and signal, and the interval containing only the masker, on the basis that threshold corresponded to a fixed signal-to-noise ratio (SNR). It has been also suggested that there is some inherent variability in one or more stages of the temporal-processing model that imposes limits on the accuracy of their operations (Shailer & Moore, 1987). Thus, for auditory tasks that do not use maskers, the concept of a decision device with a fixed SNR may still be useful if one considers the 'noise' as reflecting any inherent variability (*i.e.*, an internal 'noise' floor) that might be present in this stage. Accordingly, for modulation-detection tasks, a possible decision rule would be based on the variance of the output of the low-pass filter, as suggested by Viemeister (1979), and the assumption that threshold of modulation detection corresponds to the fixed SNR of the decision-device stage, where the 'noise' reflects the inherent variability of this stage. As the variance of a modulated sound will be greater than that of an unmodulated sound, detection of modulation in a two-alternative forced-choice task could be based on identifying the interval with the greatest variance; the threshold of modulation detection would then correspond to the modulation depth which produces a level of variance that significantly exceeds the inherent variability of this stage. An increase in the level of inherent variability in the decision device would require a concomitant increase in the variance of the output of the low-pass filter (third) stage (*i.e.*, a higher modulation-depth) to attain the fixed SNR of the decision device required for the detection of modulation.

Thus, increased levels of inherent variability in this final stage of temporal processing could in principle account for the elevated modulation-detection thresholds observed in our ASD children. Increased background neural activity in the absence of a signal (*i.e.*, increased levels of internal neural noise) is believed to be a feature of immature sensory processing, and has previously been proposed to account for reduced levels of performance of typically developing children on some auditory tasks (Nozza, 1995; Oxenham & Buus, 2000). Similarly, differential auditory system development in ASD children (*e.g.*, Gage, Siegel, & Roberts, 2003; Rapin & Katzman, 1998) is believed to reflect immature and/or delayed development of auditory sensory processing pathways, perhaps reflecting abnormalities in neural maturational processes (*i.e.*, synaptogenesis, dendritic arborization and pruning, myelination). One hypothesis is thus that abnormalities in neural maturation would result in increased levels of internal neural noise in ASD individuals, which might influence the inherent variability of the decision device.

Alternatively, reports of neurones in the IC brainstem nuclei that are sensitive to the statistical properties of acoustic stimulation have been taken as evidence in support of the existence of local mechanisms of adaptation to the acoustic properties of the input (Dean, Harper, & McAlpine, 2005; Perez-Gonzalez, Malmierca, & Covey, 2005). Consequently, deficiencies in these local processes, especially during critical developmental periods, could lead to

changes in subcortical auditory nuclei that then affect coding and processing in the ascending auditory pathways. As IC neurones also receive large corticofugal projections (*e.g.*, Coleman & Clerici, 1987; Herbert, Aschoff, & Ostwald, 1991), it is possible that cortico-collicular activity might also play a role in mechanisms for modulation detection. For example, functional ablation of the auditory cortex results in reductions in the firing rate of some IC neurones (Popelar, Nwabueze-Ogbo, & Syka, 2003), but not impaired temporal-envelope resolution (Rybalko, Mazelová, & Syka, 2001). Thus, reduced or absent corticofugal modulation of the IC could be one mechanism causing a reduction in sensitivity to modulation detection, though not temporal-envelope resolution, such as observed in the current study.

Previous reports of abnormal features in the auditory brainstem response of individuals with ASD would also appear to support deficits in low-level auditory processing in this group (*e.g.*, Galbraith, 2008; Klin, 1993; Kwon, Kim, Choe, Ko, & Park, 2007; Maziade et al., 2000; Rosenhall, Nordin, Brantberg, & Gillberg, 2003; Tas et al., 2007). However, it is worth noting, especially in the context of the increased variability of performance of our ASD children, that although abnormal brainstem auditory evoked responses (BAERs) have been observed in individuals with ASD, they would appear not to be necessary for the emergence of the disorder since it is also the case that many individuals with ASD display normal responses (Klin, 1993; Rosenhall et al., 2003). It has been suggested that abnormal BAERs may be more useful as a biomarker for characterising different subgroups within ASD (for a review, see Bomba & Pang, 2004), which may also exhibit varying degrees of sensitivity to temporal modulations and speech-in-noise recognition abilities.

4.4. Development of auditory temporal processing in children

In attempting to understand why at least some individuals with ASD might have overall poorer sensitivity to temporal modulations, it is useful to describe what is known about the development of auditory temporal processing in children in general. As attempts to compare different measures of auditory temporal-processing across studies are problematic, due to differences in the tasks used, and differences in the choice of stimuli, we will limit the comparison of previous studies with our own to those that have employed psychophysical tasks that are similar to our own.

Hall and Grose (1994) measured TMTFs in children of different ages (4–5, 6–7, and 9–10 years) and adults, using a SAM narrow-band (1000-Hz wide) continuous noise, with modulation rates of 5, 20, 100, 150, and 200 Hz. Their results showed no effect of age on temporal-envelope resolution ability; however, modulation-depth thresholds were highest for the youngest group, and progressively reduced with age, until at 9–10 years, they were not significantly different from those measured for the adults. Thus, it seems that although the level of temporal-envelope resolution may be adult-like by 4 years of age, the level of efficiency of extraction of the cues used to code temporal modulations improves with age and reaches adult-like efficiency relatively late. Whether this is indicative of age-related processing-efficiency differences specific for temporal modulations, or age-related changes in processing efficiency of sounds in general is not known. However, Hall and Grose (1994) argued for the former possibility on the basis of its explanatory power in reconciling their finding of no apparent effect of age on temporal acuity in their children with a previous report of an apparent improvement in the temporal resolution abilities of children with increasing age (*e.g.*, Grose, Hall, & Gibbs, 1993). The previous finding of age-related levels of temporal acuity was based on the difference in tone thresholds in unmodulated and modulated noise; however, the threshold differences observed across age could have been the result of changes in temporal-envelope

processing efficiency and not temporal-envelope resolution. The absence of any consistent pattern between individual modulation-detection and increment-detection thresholds in our ASD children would appear to be consistent with such an argument. Reports of adult-like processing efficiency for detection of tones in unmodulated noise in children as young as 6 years of age (Soderquist, 1993) are also consistent with this interpretation. Within this context, the elevated modulation-detection thresholds of our ASD children suggest that they are inefficient at processing temporal modulations relative to their age-matched controls. The mean age of the ASD and Control groups was just over 12 years of age. By this age, one would expect sensitivity to temporal modulations to be adult-like (Hall & Grose, 1994). A comparison of the current study's modulation-depth thresholds with those of the study of Bacon and Viemeister (1985), who also used a broadband noise to measure TMTFs in normal-hearing adults, suggests that whilst our Control-group children displayed adult-like thresholds, the modulation-depth thresholds for the ASD group more closely resembled those thresholds expected of a younger age group.

Thus, a possible interpretation of our findings is that there may be a delay in the development of auditory temporal-envelope processing associated with ASD. In this context, it is worth noting that the two out of the three best-performing ASD children (ASD5 and ASD6) were at the top end of the ASD-group age range (*i.e.*, greater than 14 years of age), and that the third best-performing ASD child (ASD3), although only 10.4 years of age, also had the highest IQ score in the study cohort. This raises a number of interesting questions, including: (1) whether the deficit in temporal-envelope processing ability observed in our ASD children resolves with increasing age, and (2) whether a deficit in temporal-envelope processing efficiency can be compensated for by, for example, increased levels of sustained attention. It is not possible to adequately address either of these questions in this study; however the young ASD adults in the study of Alcántara et al. (2004) had a mean age of about 21 years, and showed deficits in speech-in-noise thresholds in temporal-modulated noise. This suggests that the apparent deficit in amplitude-modulation detection sensitivity observed in the current study in children might not be expected to entirely resolve with increasing age, and associated speech-in-noise recognition problems may be a lasting condition that requires more direct intervention strategies. Further studies of speech-in-noise perception in ASD children of various ages will be required to confirm this possibility.

4.5. Auditory temporal-envelope processing deficits and current theories of sensory processing in ASD

We turn now to addressing the issue of whether a temporal-envelope processing deficit in ASD can be accommodated within the current understanding of sensory processing features in ASD. To date, atypical sensory-related behaviours in ASD, regardless of the sensory domain, have been interpreted within the context of two frameworks: (1) The Weak Central Coherence (WCC) Theory of Frith and Happe (1994); and (2) The Enhanced Perceptual Functioning (EPF) Model (Motttron & Burack, 2001; Motttron, Dawson, Soulières, Hubert, & Burack, 2006; Samson, Motttron, Jemel, Belin, & Ciocca, 2006). Both frameworks are useful in that they lead to predictions of a differing performance pattern in ASD on behavioural sensory tasks, which is dependent on some variable. For example, the original formulation of the WCC theory hypothesised that individuals with ASD have difficulty integrating local (low-level representation) features into 'global' (high-level representations) perceptual units (Kellerman, Fan, & Gorman, 2005; Nieto Del Rincon, 2008). However, evidence of global auditory processing in ASD has been reported, which is inconsistent with the WCC Theory (Heaton, 2005; Motttron, Peretz, Belleville, & Rouleau, 1999; Motttron, Peretz,

& Menard, 2000), and has led to a more recent formulation of WCC that emphasises a tendency or cognitive style towards superior local processing, rather than a deficit *per se* at the global level (Happé & Frith, 2006). Given the relatively low-level processing demands of the auditory stimuli used in the current study, it is difficult to interpret the deficit in amplitude-modulation detection observed within the context of the WCC theory as originally formulated, implying a deficit specifically in more 'central' processes of integration.

The EPF framework predicts that increasing stimulus spectral and/or temporal complexity will lead to worse performance, as a result of deficits in integration of information between primary and association auditory cortical areas (*e.g.*, Bonnel et al., 2010; Samson et al., 2006, 2011), whilst perception of simple stimuli is either spared or enhanced in ASD individuals (for a review, see Bertone, Motttron, Jelenic, & Faubert, 2005). In the auditory domain, the inverse relation between the (increasing) level of neural complexity of a task and the level of performance on low-level perceptual tasks is posited to be as a result of the hierarchical nature of neural processing in the auditory pathway (Langner & Schreiner, 1988).⁸

The EPF theory has had some success in accounting for the pattern of both enhanced and diminished perceptual functioning evident in ASD. It is, for example, supported by reports of an enhanced ability amongst individuals with ASD to discriminate between pure tones (Bonnel et al., 2003, 2010). Similarly, the deficits found in ASD individuals in speech-in-noise perception tasks that require the recognition of spectrotemporal complex speech signals presented simultaneously in a temporally complex background noise (Alcántara et al., 2004; Groen et al., 2009) appear to support EPF predictions of diminished performance on tasks that require high-level integration of information into more complex representations. However, the deficit observed in the current study does not support the neural complexity hypothesis of sensory processing in ASD: compared to a task requiring speech recognition in a temporally modulated background noise, the detection of amplitude modulation is a relatively simpler (non-speech) task requiring less complex processing. Moreover, EPF does not appear satisfactorily to account for the observations of seemingly intact use of complex representations in the spectral domain (Alcántara et al., 2004; Groen et al., 2009). Similarly, the EPF-predicted reductions in pitch discrimination ability for stimuli with increased temporal complexity (*i.e.*, amplitude-modulated *versus* steady-state targets) and increased spectral complexity (harmonic tone complexes *versus* pure-tone targets) have not been observed (Bonnel et al., 2010). Thus, it seems neither the WCC or EPF theories, in their current instantiations, provide a comprehensive account of auditory sensory processing in ASD; nor can they fully accommodate a specific deficit in processing efficiency of temporal modulations.

4.6. Auditory processing efficiency and speech-in-noise perception deficits in ASD and language-based learning disorders

The presence of speech-in-noise perception deficits in individuals diagnosed with developmental dyslexia (DD) and specific language impairment (SLI) raises the possibility that the

⁸ The auditory system contains several levels of systematic topographical organization with respect to the response characteristics that convey temporal modulation aspects of the auditory signal. Indeed, this is often described as the most basic temporal organizational feature of auditory processing, and is reflected by the observed general reduction in the temporal activity patterns of neural elements along the auditory pathway: temporal resolution is highest at the level of the auditory nerve (up to 5000 Hz; Palmer, 1987); is reduced at the level of the IC (up to 1000 Hz; Langner & Schreiner, 1988; Schreiner & Langner, 1984); and is apparently worst at the level of the auditory cortex (about 20 Hz; Creutzfeldt, Hellweg, & Schreiner, 1980).

processing-efficiency deficit observed in the current study might not be a distinctive feature of the ASD sensory phenotype, but one which may be shared with other developmental disorders (Ziegler, Pech-Georgel, George, Alario, & Lorenzi 2005; Ziegler, Pech-Georgel, George, & Lorenzi, 2009, 2011). However, although all three developmental-disorder groups exhibit speech-in-noise perception deficits, individuals with ASD have reduced masking release (Alcántara et al., 2004; Groen et al., 2009), whereas individuals with DD and SLI appear to have normal levels of masking release (Ziegler et al., 2005, 2009, 2011). In other words, individuals with ASD appear to perform as well as controls on speech-in-noise tasks with unmodulated backgrounds, but the recognition of speech in backgrounds that contain temporal modulations is reduced relative to controls. In contrast, individuals with DD and SLI have reduced levels of speech-in-noise recognition that are very similar in magnitude in both unmodulated and modulated backgrounds.

Ziegler and colleagues have suggested that the presence of normal levels of masking release suggests that low-level auditory processing capabilities (e.g., spectral and temporal resolution) of their SLI and DD groups are no different to that of controls, and therefore that the speech-in-noise deficits observed are the result of *reduced efficiency* of higher-level linguistic-related auditory processes, which are only revealed when the redundancy of the speech signal is significantly reduced as a result of the presence of background noise (Ziegler et al., 2005, 2009). Specifically, the speech-in-noise deficits observed in DD and SLI have been interpreted as central auditory deficits in the mapping of acoustic cues onto phonetic and phonological categories (Ziegler et al., 2005, 2009).

It is apparent, however, that any direct comparison of the speech-in-noise deficits observed in ASD and those reported in DD and SLI, which is based on processing-efficiency deficits, is potentially limited by the way in which “auditory-processing efficiency” is defined across different studies. Accordingly, it is more informative to examine the extent to which the speech-in-noise deficits reported by Ziegler and colleagues for DD and SLI might be related to those deficits observed in ASD (Alcántara et al., 2004; Groen et al., 2009), when a more narrow definition of auditory-processing efficiency is applied to all of the groups; namely, auditory-processing efficiency as operationalised within the framework of the temporal-processing model shown in Fig. 7. It is notable that a very similar temporal-processing model provides a good account of the tone-in-noise threshold deficits in SLI individuals, under simultaneous, forward, and backward masking conditions, in terms of reduced levels of processing efficiency as a result of a reduced signal-to-noise ratio in the final (decision-device) stage of the model (Hartley & Moore, 2002). One might also expect that reduced levels of processing efficiency would be successful in accounting for the elevated modulation-detection thresholds that have been previously observed in DD (Menell et al., 1999; Lorenzi et al., 2000; Rocheron et al., 2002). Moreover, the presence of reduced modulation-detection abilities in DD, but no deficit in detection of a tone in simultaneously presented unmodulated noise (e.g., Amitay, Ahissar, & Nelken, 2002; McAnally & Stein, 1996; Rosen & Manganari, 2001; Wright et al., 1997), is consistent with a relatively low-level auditory processing-efficiency deficit for temporal modulations in DD. The extent to which this might also be true in SLI is difficult to determine currently. Individuals with SLI do appear to have higher tone-in-noise thresholds in simultaneous-masking tasks relative to controls (McAnally & Stein, 1996; Rosen, Adlard, & van der Lely, 2009; Wright et al., 1997), consistent with the presence of a general processing-efficiency deficit in this group. However, there are no studies of the modulation-detection abilities of SLI individuals, so more cannot yet be said about their temporal processing efficiency.

Thus, there is some evidence from DD and SLI studies that have used non-speech stimuli, which is consistent with the presence of a low-level auditory processing-efficiency deficit in these groups. However, the extent to which this narrow definition of auditory-processing efficiency can account for levels of performance on a speech-in-noise task is clearly limited. In contrast to non-verbal auditory stimuli, the acoustic cues important for decoding the speech signal are multidimensional in nature: there may be several different acoustic cues for a given phoneme, of which only one or two might be sufficient for decoding purposes. This allows for a high level of redundancy in the speech signal, which might be helpful in maintaining good speech intelligibility in adverse listening conditions, when the number of available cues is reduced. Moreover, speech perception involves processing stages at many different levels and most likely requires simultaneous integration of acoustic cues across different time scales to provide robust representations of the linguistic information available within the speech signal (Greenberg, 1998). Using a similar argument to that proposed by Ziegler et al. (2009), the presence of external noise might result in selective reduction of those acoustic cues most vulnerable to the corrupting influence of noise. Thus, a speech-in-noise deficit might reflect a reduction in the amount of information available from the integration of different acoustic cues, and provide less-reliable or relatively ambiguous information for mapping acoustic cues onto phonetic and/or phonological categories.

Even when speech stimuli are used across different studies, the pattern of speech-in-noise deficits is likely to be influenced by differences in psycholinguistic processing demands. For example, as previously noted, the pattern of thresholds for tone-in-noise and modulation-detection tasks in DD are consistent with a temporal processing-efficiency deficit in this group, similar to that found for our ASD children. If this were the case one might have expected speech masking-release levels to be reduced in the study of Ziegler et al. (2009). However, a difference between the speech-in-noise studies in ASD and DD and SLI is that whilst the former employed sentences with a high element of predictability, the latter two used vowel-consonant-vowel (VCV) nonsense-syllable stimuli. The two types of speech stimuli target different levels of psycholinguistic processing, which might also account for the different levels of masking release observed, particularly given that the use of linguistic content is believed to be weak in ASD.

Despite the limitations of a temporal-processing model in providing a comprehensive account of differences in performance on speech-in-noise tasks in ASD, and also DD and SLI, the reported auditory deficits in DD and SLI on tasks that have employed non-speech stimuli appear to be consistent with a low-level auditory processing-efficiency deficit in these groups. Moreover, as there is no compelling reason to believe that sub-cortical processing of a speech signal is qualitatively different to that of non-speech stimuli, it is possible that a lower-level auditory processing-efficiency deficit may also contribute to measures of speech-in-noise recognition in these groups. Whether this is indeed the case is not possible to determine presently. Direct comparisons of the speech-in-noise deficits in ASD and other developmental disorders that also exhibit poor speech-in-noise perception (e.g., DD and SLI) will not be fully informative until there is a better understanding of the contribution of any low-level auditory processing-efficiency deficits on speech-in-noise performance. To this end, contemporaneous measures of speech-in-noise recognition, thresholds of detection of tones in unmodulated and modulated noise, and modulation-detection thresholds, for individuals with DD, SLI, and ASD, will be required before the speech-in-noise deficits present in these different groups can be appropriately compared and quantitatively related to any non-speech low-level auditory-processing deficits that may also be present.

More generally, in attempting to relate the contribution of any observable non-verbal auditory deficits to communication impairments observed in developmental disorders, it would be informative to include non-verbal auditory stimuli that have been constructed to contain speech-like acoustic characteristics (e.g., Goswami et al., 2002). This is an important issue as symptoms such as reading deficits, language difficulties, inattention, speech-in-noise perception deficits, and non-verbal auditory-processing deficits – including reduced auditory-processing efficiency – can co-occur in several language-based learning disorders (e.g., DD and SLI) and ASD, and also within individuals comprising any one disorder, albeit with varying frequency and severity. Given the possible consequences of auditory deficits on language and literacy learning difficulties to school performance the nature of the relationship between supra-threshold non-verbal auditory-processing abilities and communication skills in a variety of populations is of more than academic interest. This was recently highlighted by a report on the need for improved and more reliable diagnoses of central auditory processing disorder (CAPD), which also features speech-in-noise perception deficits amongst an array of reported deficits in performance on non-speech auditory tasks (BSA, 2011).

4.7. Auditory-processing deficits and language impairment in ASD and language-based learning disorders

The interaction between auditory sensory deficits (and apparent enhancements in pitch discrimination abilities) and language and communication impairments that characterise the ASD phenotype is largely unknown at the present time. In contrast, the co-occurrence of auditory deficits (e.g., forward and backward masking tone-in-noise thresholds) and language and literacy problems has led some researchers to claim that the presence of low-level non-verbal auditory deficits in individuals with DD and SLI is a good predictor of the development of their language and communications skills (e.g., Tallal, Miller, Jenkins, & Merzenich, 1997). However, the veracity of this claim has recently been seriously called into question by a number of researchers (e.g., Rosen et al., 2009). Rosen (2003) argues that as only some but not all auditory skills are impaired in individuals with SLI and DD, and there is little or no relationship between the severity of the auditory deficits and language disorders in these groups, there is therefore no strong evidence to support the claim that auditory deficits lead to language disorders in these populations. A more parsimonious interpretation of the co-occurrence of auditory and language deficits in DD and SLI is that one simply occurs in association with the other. Similarly, the presenting symptoms of CAPD were found to be largely unrelated to auditory sensory processing. Rather, response variability and cognitive performance were the best predictors of poor communication and listening, suggesting that CAPD is primarily an attentional problem (Moore, Ferguson, Edmondson-Jones, Ratib, & Riley, 2010).

In light of previous assertions of a link between auditory deficits and language development in DD and SLI, the presence of speech-in-noise deficits in ASD (Alcántara et al., 2004; Groen et al., 2009), and the low-level auditory-processing deficit found in our ASD children, one might also question whether an auditory processing-efficiency deficit in ASD might impact on the development of language in ASD. Two questions are of particular relevance: (1) might speech-in-noise deficits be related to language delay/impairment in ASD, and (2) might non-verbal low-level auditory-processing deficits, such as reduced temporal-processing efficiency, be related to language delay/impairment in ASD?

Assuming Verbal IQ measures can be taken as a good predictor of receptive language abilities, it would appear then that the within-normal-range Verbal IQ scores, but impaired speech-in-noise abilities, of those ASD individuals who took part in the

studies of Alcántara et al. (2004) and Groen et al. (2009), would argue against a significant contribution of speech-in-noise deficits on the receptive language abilities of individuals with ASD, in at least those with normal-range IQ. However, the relation between speech-in-noise and processing-efficiency deficits and apparently normal receptive language abilities in at least some ASD individuals does not do justice to the complexity of the language process; nor does it encompass the larger language phenotype in ASD. For example, aspects of speech recognition related to semantic, syntax, and circumstantial cues are of considerable importance, especially in noisy environments where the redundancy of information present in the speech signal is significantly reduced. Impairment in pragmatic (socially contextualised) language and communication is an accepted hallmark of all ASDs whatever the general cognitive level and includes difficulty in understanding the meaning of words, grammar, syntax, prosody, eye gaze, body language, gestures, or social context. Whilst it is highly unlikely that such a broad array of communication difficulties might be a direct and simple result of sensory processing deficits, an auditory temporal processing-efficiency deficit could reduce the salience of supra-segmental speech information, cued by the fluctuations in amplitude over time of the speech waveform, and thus might account for the reduced receptive prosodic skills that have been observed in children with ASD (McCann, Peppé, Gibbon, O'Hare, & Rutherford, 2007; Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2007).

There is also good evidence that the temporal envelope of the speech waveform also provides information about the phonemic structure of speech. For example, the amplitude envelope of the speech waveform provides important acoustic cues primarily for manner of articulation contrasts (Rosen, 1992), though it has also been shown to contribute to the discrimination of voiced from voiceless speech sounds; for indicating the amplitude of the consonant relative to the vowel sound; and for discriminating the place of articulation of voiceless plosive consonants (van Tassel, Soli, Kirby, & Widin, 1987). Thus, an auditory processing-efficiency deficit, whether it apply to sounds in general, or specifically to temporal modulations of sounds, could have a potential impact on the formation or perceptual fine tuning of phonetic and/or phonological categories, which have clear implications for the development of language abilities. A relatively small difference in auditory processing developing prenatally could thus have long-term and ultimately significant developmental consequences.

Deficits in the transmission of features of consonant voicing in children with SLI (Ziegler et al., 2005), and consonant voicing, manner of articulation, and place of articulation in children with DD (Ziegler et al., 2009), have been reported, although the extent to which these deficits are the result of low-level processing-efficiency deficits is not known at the present time. Unfortunately, there is currently no information on the transmission of consonant-feature information in quiet or in amplitude-modulated noise (with a range of modulation rates) in children with ASD. Such measures would be informative for a better understanding of the reduced masking-release levels observed in ASD, by identifying the nature of the acoustic cues of speech that are not transmitted well and therefore result in greater confusions for the listener. Moreover, a better understanding of the nature of any speech-perception deficits (e.g., greater consonant confusions) that can be confidently attributed to the presence of low-level auditory-processing deficits (e.g., processing-efficiency of temporal modulations) would improve our understanding of the relation between auditory processing and receptive-language abilities in ASD, and also the role that non-auditory factors, such as attention and cognitive abilities, might play in the performance level observed for ASD individuals on speech and non-speech auditory-perception tasks.

5. Summary

Auditory temporal-envelope processing in 6 children with ASD, as measured using the detection of amplitude modulation as a function of its modulation rate (*i.e.*, the Temporal Modulation Transfer Function) appears to be abnormal when compared to performance of age- and IQ-matched typically developing controls on the same task. Specifically, the modulation-depth thresholds for the ASD children were significantly elevated, suggesting a reduction in sensitivity to temporal modulations, regardless of the rate of the modulations. Accordingly, our findings lend support to the temporal processing deficit hypothesis proposed by Alcántara et al. (2004) to account for the reduced masking-release levels for temporally modulated backgrounds previously observed for individuals with ASD.

A plausible interpretation of our findings is that the children in the ASD group were less efficient at extracting those cues that code for temporal modulation, although as a group, they appeared to have no greater difficulty than typically developing children in extracting those cues that code for overall level. Given the relatively small sample sizes employed in the current study, and the presence of greater variability in performance in the ASD group, further studies will be required before the findings of the current study can be generalised to the wider ASD population. Furthermore, as the 6 children with ASD who participated in this study all had a diagnosis of Asperger's syndrome, our findings cannot be necessarily generalised to other sub-groups within the ASD domain (*e.g.*, high-functioning individuals with autism).

It is hypothesised that the reduced temporal-processing efficiency in the ASD group is a consequence of a less-sensitive higher-level decision device. The results are consistent with abnormalities in processing in the inferior colliculi nuclei of the auditory brainstem and/or increased levels of internal neural noise. Whether this is a consequence of abnormal corticofugal modulation of peripheral auditory processing centres, and/or deficiencies in local processes of neural adaptation during critical developmental periods that lead to changes in sub-cortical structures and coding and processing in the ascending auditory pathways, is at present not known.

Role of the funding source

The Nancy Lurie Marks Family Foundation played no part in the study design, the collection, analysis and interpretation of the data, in the writing of this report or the decision to submit the paper for publication.

Conflicts of interests

There are no actual or potential conflicts of interest for the authors.

Acknowledgments

This study would not have been possible without the interest and generous support of the Nancy Lurie Marks Family Foundation (Research project grant, 20453). We thank Christian Füllgrabe, Brian Moore, and Kate Plaisted-Grant, for helpful comments on previous versions of this paper. We also owe a debt of gratitude to the two anonymous reviewers for their valuable and constructive comments that significantly improved the clarity of this paper. Finally, though not least, we are very grateful to the children who took part in the study, and to their parents for their willingness to donate their time.

References

- Alcántara, J. I., Holube, I., & Moore, B. C. J. (1996). Effects of phase and level on vowel identification: Data and predictions based on a nonlinear basilar-membrane model. *Journal of the Acoustical Society of America*, 100, 2382–2392.
- Alcántara, J. I., & Moore, B. C. J. (1995). The identification of vowel-like harmonic complexes: Effects of component phase, level, and fundamental frequency. *Journal of the Acoustical Society of America*, 97, 3813–3824.
- Alcántara, J. I., Weisblatt, E. J., Moore, B. C. J., & Bolton, P. F. (2004). Speech perception in high-functioning subjects with autism or Asperger's syndrome. *Journal of Child Psychiatry and Psychology*, 45, 1107–1114.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed., Rev). Washington, DC.
- Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory processing deficits in reading disabled adults. *Journal of the Association for Research in Otolaryngology*, 3, 302–320.
- Bacon, S. P., & Viemeister, N. F. (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects. *Audiology*, 24, 117–134.
- Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain*, 128, 2430–2441.
- Bishop, D., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment. *Journal of Speech, Language, and Hearing Research*, 42, 1295–1310.
- Bomba, M. D., & Pang, E. W. (2004). Cortical auditory evoked potentials in autism: A review. *International Journal of Psychophysiology*, 53, 161–169.
- Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., et al. (2010). Enhanced pure-tone discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia*, 48, 2465–2475.
- Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A.-M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience*, 15, 226–235.
- Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
- British Society of Audiology. (2004). *Recommended Procedure for pure tone air and bone conduction threshold audiometry with and without the use of masking and determination of uncomfortable loudness levels*. www.thebsa.org.uk
- British Society of Audiology. (2011). *Position statement: Auditory processing disorder*. www.thebsa.org.uk
- Coleman, J. R., & Clerici, W. J. (1987). Sources of projections to subdivisions of the inferior colliculus in the rat. *Journal of Comparative Neurology*, 262, 215–226.
- Cooke, M. (2006). A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America*, 119, 1562–1573.
- Creutzfeldt, O. D., Hellweg, F. C., & Schreiner, C. E. (1980). Thalamo-cortical transformation of responses to complex auditory stimuli. *Experimental Brain Research*, 39, 87–104.
- Dahlgren, S. O., & Gillberg, C. (1989). Symptoms in the first two years of Life. A preliminary population study of infantile autism. *European Archives of Psychiatry and Neurological Sciences*, 238, 169–174.
- Dean, I., Harper, N. S., & McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, 8, 1684–1689.
- Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *Journal of the Acoustical Society of America*, 95, 1053–1064.
- Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *Journal of the Acoustical Society of America*, 74, 739–743.
- Eddins, D. A. (1993). Absolute modulation detection of narrow-band noise: Effects of amplitude bandwidth and frequency region. *Journal of the Acoustical Society of America*, 93, 470–479.
- Festen, J., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, 88, 1725–1736.
- Fletcher, H. (1940). Auditory patterns. *Reviews in Modern Physics*, 12, 47–66.
- Forrest, T. G., & Green, D. M. (1987). Detection of partially filled gaps in noise and the temporal modulation transfer function. *Journal of the Acoustical Society of America*, 82, 1933–1943.
- Frisina, R. D. (2001). Subcortical neural coding mechanisms for auditory temporal processing. *Hearing Research*, 158, 1–27.
- Frith, U., & Happe, F. (1994). Autism: Beyond theory of mind. *Cognition*, 50, 115–132.
- Füllgrabe, C., Berthommier, F., & Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hearing Research*, 211, 74–84.
- Gage, N. M., Siegel, B., & Roberts, T. P. L. (2003). Cortical auditory system maturational abnormalities in children with autism spectrum disorder: An MEG investigation. *Developmental Brain Research*, 144, 201–209.
- Galbraith, G. C. (2008). Editorial: Deficient brainstem encoding in autism. *Clinical Neurophysiology*, 119, 1697–1700.
- George, E. L. J., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 120, 2295–2311.
- Gnansia, D., Jourdes, V., & Lorenzi, C. (2008). Effect of masker modulation depth on speech masking release. *Hearing Research*, 239, 60–68.
- Goldfarb, W. (1963). Self-awareness in schizophrenic children. *Archives of General Psychiatry*, 81, 47–68.
- Goswami, U., Thomson, J., Richardson, U., Stinkhorn, R., Hughes, D., Rosen, S., et al. (2002). Amplitude envelope onsets and developmental dyslexia: A new

- hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 10911–10916.
- Grandin, T., & Scariano, M. M. (1986). *Emergence: Labelled autistic*. Costello: Tunbridge Wells.
- Gravel, J. S., Dunn, M., Lee, W. W., & Ellis, M. A. (2006). Peripheral audition of children on the autistic spectrum. *Ear & Hearing*, 27, 299–312.
- Greenberg, S. (1998). Recognition in a new key – Towards a science of spoken language. In *Proceedings of the international conference on acoustics, speech, and signal processing* Seattle, (pp. 1041–1044).
- Groen, W. B., van Orzo, L., tar Horne, N., Winkles, S., van der Gaga, R.-J., Butler, J. K., et al. (2009). Intact spectral but abnormal temporal processing in autism. *Journal of Autism and Developmental Disorders*, 39, 742–750.
- Grose, J. H., Hall, J. W., III, & Gibbs, C. (1993). Temporal analysis in children. *Journal of Speech & Hearing Research*, 36, 351–356.
- Gustafsson, H. A., & Arlinger, S. D. (1994). Masking of speech by amplitude-modulated noise. *Journal of the Acoustical Society of America*, 95, 518–529.
- Hall, J. W., III, & Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *Journal of the Acoustical Society of America*, 96, 150–154.
- Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36, 5–25.
- Hartley, D. E. H., Hill, P. R., & Moore, D. R. (2003). The auditory basis of language impairments: Temporal processing versus processing efficiency hypothesis. *International Journal of Paediatric Otorhinolaryngology*, 67S1, S137–S142.
- Hartley, D. E. H., & Moore, D. R. (2002). Auditory processing efficiency deficits in children with developmental language impairments. *Journal of the Acoustical Society of America*, 112, 2962–2966.
- Hartley, D. E., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech Language and Hearing Research*, 43, 1402–1415.
- Heaton, P. (2003). Pitch memory, labeling, and disembodiment in autism. *Journal of Child Psychology and Psychiatry*, 44, 543–551.
- Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders*, 35, 787–793.
- Herbert, H., Aschoff, A., & Ostwald, J. (1991). Topography of projections from the auditory cortex to the inferior colliculus in the rat. *Journal of Comparative Neurology*, 304, 103–122.
- Hermelin, B., & O'Connor, N. (1970). *Psychological experiments with autistic children*. Oxford, England: Paragon.
- Hill, P. R., Hartley, D. E. H., Glasberg, B. R., Moore, B. C. J., & Moore, D. R. (2004). Auditory processing efficiency and temporal resolution in children and adults. *Journal of Speech, Language, and Hearing Research*, 47, 1022–1029.
- Howard-Jones, P. A., & Rosen, S. (1993). The perception of speech in fluctuating noise. *Acoustical*, 78, 258–272.
- Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders*, 36, 77–90.
- Jones, C. R. G., Happé, F., Baird, G., Simon off, E., Mars den, A. J. S., Treaty, J., et al. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia*, 47, 2850–2858.
- Joris, P. X., & Yin, T. C. T. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *Journal of the Acoustical Society of America*, 91, 215–232.
- Joris, P. X., & Yin, T. C. T. (1998). Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *Journal of Neurophysiology*, 79, 253–269.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.
- Kellerman, G. R., Fan, J., & Gorman, J. M. (2005). Auditory abnormalities in autism: Toward functional distinctions among findings. *CNS Spectrums*, 10, 748–756.
- Klin, A. (1993). Auditory brainstem responses in autism: Brainstem dysfunction or peripheral hearing loss? *Journal of Autism and Developmental Disorders*, 23, 15–35.
- Kwon, S., Kim, J., Choe, B.-H., Ko, C., & Park, S. (2007). Electrophysiologic assessment of central auditory processing by auditory brainstem responses in children with autism spectrum disorders. *Journal of Korean Medical Science*, 22, 656–659.
- Kwon, B. J., & Turner, C. W. (2001). Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference? *Journal of the Acoustical Society of America*, 110, 1130–1140.
- Langner, G., & Schreiner, C. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanism. *Journal of Neurophysiology*, 60, 1799–1822.
- Leekam, S. R., Libby, S. J., Wing, L., Gould, J., & Taylor, C. (2002). The diagnostic interview for social and communication disorders: Algorithms for ICD-10 childhood autism and Wing and Gould autistic spectrum. *Journal of Child Psychiatry and Psychology*, 43, 327–342.
- Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism and Developmental Disorders*, 37, 894–910.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49, 467–477.
- Liang, L., Lu, T., & Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *Journal of Neurophysiology*, 87, 2237–2261.
- Lorenzi, C., Dumont, A., & Füllgrabe, C. (2000). Use of temporal envelope cues by children with Dyslexia. *Journal of Speech, Language, and Hearing Research*, 43, 1367–1379.
- Lorenzi, C., Husson, M., Adroit, M., & Debruille, X. (2006). Speech masking release in listeners with flat hearing loss: Effects of masker fluctuation rate on identification scores and phonetic feature reception. *International Journal of Audiology*, 45, 487–495.
- Macleod, A., & Summerfield, A. Q. (1990). A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British Journal of Audiology*, 24, 29–44.
- Maziade, M., Merette, C., Cayer, M., Roy, M. A., Szatmari, P., & Cote, R. (2000). Prolongation of brainstem auditory-evoked responses in autistic probands and their unaffected relatives. *Archives of General Psychiatry*, 57, 1077–1083.
- McAnally, K. I., & Stein, J. F. (1996). Auditory temporal coding in dyslexia. *Proceedings of the Royal Society of London Series B – Biological Sciences*, 263, 961–965.
- McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language & Communication Disorders*, 42, 682–702.
- Menell, P., McAnally, K. I., & Stein, J. F. (1999). Psychophysical sensitivity and physiological response to amplitude modulation in adult dyslexic listeners. *Journal of Speech, Language, and Hearing Research*, 42, 797–803.
- Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, 271, 77–81.
- Miller, G. A. (1947). The masking of speech. *Psychological Bulletin*, 44, 105–129.
- Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 22, 167–173.
- Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). London: Academic Press.
- Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., & Riley, A. (2010). Nature of auditory processing disorder in children. *Pediatrics*, 126, e382–e390.
- Moore, B. C. J., Glasberg, B. R., Plack, C. J., & Biswas, A. K. (1988). The shape of the ear's temporal window. *Journal of the Acoustical Society of America*, 83, 1102–1116.
- Mottron, L., & Burack, J. (2001). Enhanced perceptual functioning in the development of persons with autism. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 131–148). Hillsdale, NJ: Erlbaum.
- Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36, 27–43.
- Mottron, L., Peretz, I., Belleville, S., & Rouleau, N. (1999). Absolute pitch in autism: A case study. *Neurocase*, 5, 485–501.
- Mottron, L., Peretz, I., & Menard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 1057–1065.
- Nelson, P. B., Jin, S.-H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal hearing listeners. *Journal of the Acoustical Society of America*, 113, 961–968.
- Nieto Del Rincon, P. L. (2008). Autism: Alterations in auditory perception. *Reviews in the Neurosciences*, 19, 61–78.
- Nozza, R. J. (1995). Estimating the contribution of non-sensory factors to infant-adult differences in behavioral thresholds. *Hearing Research*, 91, 72–78.
- Ornitz, E. M. (1974). The modulation of sensory input and motor output in autistic children. *Journal of Autism & Childhood Schizophrenia*, 7, 207–229.
- Oxenham, A. J., & Buus, S. (2000). Level discrimination of sinusoids as a function of duration and level for fixed-level, roving-level, and across-frequency conditions. *Journal of the Acoustical Society of America*, 107, 1605–1614.
- Oxenham, A., & Moore, B. C. J. (1994). Modeling the additivity of non-simultaneous masking. *Hearing Research*, 80, 105–118.
- Palmer, A. R. (1987). Physiology of the cochlear nerve and cochlear nucleus. In M. P. Haggard, & E. F. Evans (Eds.), *Hearing*. Edinburgh: Churchill Livingstone.
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., & Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *Journal of the Acoustical Society of America*, 72, 1788–1803.
- Penner, M. J. (1977). Detection of temporal gaps in noise as a measure of the decay of the auditory sensation. *Journal of the Acoustical Society of America*, 61, 552–557.
- Peppé, S., McCann, J., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research*, 50, 1015–1028.
- Perez-Gonzalez, D., Malmierca, M. S., & Covey, E. (2005). Novelty detector neurons in the mammalian auditory midbrain. *European Journal of Neuroscience*, 22, 2879–2885.
- Plaisted, K. C., Saksida, L., Alcántara, J. I., & Weisblatt, E. J. (2003). Towards an understanding of the mechanisms of weak central coherence: Experiments in configural learning and auditory perception. *Philosophical Transactions: Biological Sciences*, 358, 375–386.
- Popelar, J., Nwabueze-Ogbo, F. C., & Syka, J. (2003). Changes in neuronal activity of the inferior colliculus in rat after temporal inactivation of the auditory cortex. *Physiological Research*, 52, 615–628.
- Psychological Corporation. (1999). *Wechsler abbreviated scale of intelligence manual*. San Antonio, TX: Psychological Corporation.
- Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *Journal of the Acoustical Society of America*, 114, 446–454.
- Rapin, I., & Katzman, R. (1998). Neurobiology of autism. *Annals of Neurology*, 43, 7–14.
- Rocheron, I., Lorenzi, C., Füllgrabe, C., & Dumont, A. (2002). Temporal envelope processing in dyslexic children. *NeuroReport*, 13, 1683–1687.

- Rodenburg, M. (1977). Investigation of temporal effects with amplitude modulated signals. In E. F. Evans, & J. P. Wilson (Eds.), *Psychophysics and physiology of hearing* (pp. 429–437). London: Academic Press.
- Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions: Biological Sciences*, 336, 367–373.
- Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics*, 31, 509–527.
- Rosen, S., Adlard, A., & van der Lely, H. K. J. (2009). Backward and simultaneous masking in children with grammatical specific language impairment: No simple link between auditory processing in children with dyslexia? *Journal of Speech, Language, and Hearing Research*, 52, 396–411.
- Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia? *Journal of Speech, Language, and Hearing Research*, 44, 720–736.
- Rosenhall, U., Nordin, V., Brantberg, K., & Gillberg, C. (2003). Autism and auditory brain stem responses. *Ear and Hearing*, 24, 206–214.
- Rosenhall, U., Nordin, V., Sandstrom, M., Ahlsen, G., & Gillberg, C. J. (1999). Autism and hearing loss. *Journal of Autism and Developmental Disorders*, 29, 349–357.
- Rybalko, N., Mazelová, S., & Syka, J. (2001). Influence of auditory cortex ablation and noise exposure on gap detection thresholds in rats. In *Abstract of the Twenty-fourth ARO Meeting* St. Petersburg Beach, 4–8. 2.2001. No. 305, (p. 86).
- Samson, F., Hyde, K. L., Bertone, A., Soulieres, I., Mendrek, A., Ahad, P., et al. (2011). Atypical processing of auditory temporal complexity in autistics. *Neuropsychologia*, 49, 546–555.
- Samson, F., Mottron, L., Jemel, B., Belin, P., & Ciocca, V. (2006). Can spectro-temporal complexity explain the autistic pattern of performance on auditory tasks? *Journal of Autism and Developmental Disorders*, 36, 65–76.
- Schreiner, C., & Langner, G. (1984). Representation of periodicity information in the inferior colliculus of the cat. *Society of Neuroscience Abstracts*, 10, 395.
- Schweizer, K., & Moosbrugger, H. (2004). Attention and working memory as predictors of intelligence. *Intelligence*, 32, 329–347.
- Sellick, P. M., Patuzzi, R., & Johnstone, B. M. (1982). Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *Journal of the Acoustical Society of America*, 61, 133–149.
- Shailer, M. J., & Moore, B. C. J. (1987). Gap detection and the auditory filter: Phase effects using sinusoidal stimuli. *Journal of the Acoustical Society of America*, 81, 1110–1117.
- Soderquist, D. R. (1993). Auditory filter widths in children and adults. *Journal of Experimental Child Psychology*, 56, 371–384.
- Summers, V., & Molis, M. R. (2004). Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level. *Journal of Speech, Language, and Hearing Research*, 47, 245–256.
- Sutcliffe, P., Bishop, D. V. M., Houghton, S., & Taylor, M. (2006). Effect of attentional state on frequency discrimination: A comparison of children with ADHD on and off medication. *Journal of Speech, Language, and Hearing Research*, 49, 1072–1084.
- Takahashi, G. A., & Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech and Hearing Research*, 35, 1410–1421.
- Talbot, J. B., Witton, C., Hebb, G. S., Stoodley, C. J., Westwood, E. A., France, S. J., et al. (2002). On the relationship between dynamic visual and auditory processing and literacy skills: Results from a large primary-school study. *Dyslexia*, 8, 204–225.
- Tallal, P. (1984). Temporal or phonetic processing deficit in dyslexia? That is the question. *Applied Psycholinguistics*, 5, 167–169.
- Tallal, P., Miller, S. L., Jenkins, W. M., & Merzenich, M. M. (1997). The role of temporal processing in developmental language-based learning disorders. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 49–66). Mahwah, NJ: Erlbaum.
- Tas, A., Yagiz, R., Tas, M., Esme, M., Uzun, C., & Karasalioglu, A. R. (2007). Evaluation of hearing in children with autism by using TEOAE and ABR. *Autism*, 11, 73–79.
- van Tassel, D. J., Soli, S. D., Kirby, V. M., & Widin, G. P. (1987). Speech waveform cues for consonant recognition. *Journal of the Acoustical Society of America*, 82, 1152–1161.
- Viemeister, N. F. (1979). Temporal modulation transfer functions based on modulation threshold. *Journal of the Acoustical Society of America*, 66, 1364–1380.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392–393.
- Werner, L. A., & Bargones, J. Y. (1991). Sources of auditory masking in infants: Distraction effects. *Perception & Psychophysics*, 50, 405–412.
- Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature*, 387, 176–178.
- Young, E. D. (2008). Neural representations of spectral and temporal information in speech. *Philosophical Transactions of the Royal Society B*, 363, 923–945.
- Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F.-X., & Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 14110–14115.
- Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, 12, 732–745.
- Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2011). Noise on, voicing off: Speech perception deficits in children with specific language impairment. *Journal of Experimental Child Psychology*, 110, 362–372.