Taylor & Francis
health sciences

Prosody in autism spectrum disorders: a critical review

Joanne McCann and Sue Peppé

Queen Margaret University College, Edinburgh, UK

Abstract

Background: Many individuals with autism spectrum disorders present with unusual or odd-sounding prosody. Despite this widely noted observation, prosodic ability in autism spectrum disorders is often perceived as an under-researched area.

Aims: This review seeks to establish whether there is a prosodic disorder in autism, what generalizations can be made about its various manifestations and whether these manifestations vary according to the diagnosis. A literature review was carried out to establish what areas of prosody in autism spectrum disorders have been researched to date, what the findings have been and to determine what areas are yet to be researched.

Main contribution: It is shown that prosody in autism spectrum disorders is an under-researched area and that where research has been undertaken, findings often conflict. The findings of these conflicting studies are compared and recommendations are made for areas of future research.

Conclusions: Research in this area has covered mostly prosodic expression, although some more recent studies cover comprehension, processing and the relationship of receptive prosodic ability to theory of mind. Findings conflict and methodology varies greatly.

Keywords: prosody, intonation, autism, Asperger's syndrome.

Introduction

Autism

Autism is a pervasive developmental disorder (PDD) and has been defined as a triad of impairment: atypical development in reciprocal social interaction; atypical communication; and restricted, stereotyped and repetitive behaviours (Wing and

Address correspondence to: Joanne McCann, Queen Margaret University College, Clerwood Terrace, Edinburgh EH12 8TS, UK; e-mail: jmccann@qmuc.ac.uk

Gould 1979, Aitken *et al.* 1998). It is a disorder that begins in the first 36 months of life (DSM-IV 1994) and social impairment is now seen by many as the primary symptom (Baron-Cohen 1989). Autism is a spectrum disorder ranging from low-functioning autism (individuals may be non-verbal) with associated learning difficulties to high-functioning autism (HFA) and Asperger's syndrome (AS). The distinction between HFA and AS is controversial. AS is also a PDD with impaired social interaction and repetitive, restricted and stereotyped behaviours, but the most recent diagnostic criteria holds that individuals with AS do not demonstrate a general language delay (DSM-IV 1994). This review deals principally with autism, or more often HFA, and also with AS.

Prosody

As a feature of impaired communication in autism, individuals often display disordered prosody (Baltaxe and Simmons 1985), but it is unclear from the literature what is meant by this. Prosody plays an important role in a range of communicative functions (affective, pragmatic, syntactic), serving to enhance or change the meaning of what is said (Couper-Kuhlen 1986). In the literature, the prosodic terms used refer variously to prosodic function and/or to its exponency. For example, the term 'stress' can refer to the conveying of focus on a linguistic element and also to the exponency or means by which this is done (variation in prosodic elements such as pitch/fundamental frequency, loudness/intensity, duration, pause/silence, speech rate and rhythm; Crystal 1969, Quirk *et al.* 1985). The distinction is an important one because prosodic elements have different roles, paralinguistic as well as linguistic, and disorder may be at the 'form' level, e.g. a lack of ability to perceive difference between pitch levels, or at the 'function' level, e.g. an inability to appreciate the communicative role of stress.

 This review has taken a mainly functional approach, i.e. considering the various functions of prosody and whether individuals with autism use them effectively, since this is the approach taken by the majority of studies. Aspects of speech with communicative effect but which are linguistically unquantifiable, such as vocal quality and articulatory settings, are not included. Similarly, the indexical functions of prosody (such as regional accents) are not considered, only functions that can be varied by individuals, more or less intentionally, for communicative effect.

 Communicative functions in which prosodic elements play an important part include the following:

- Pragmatic functions such as:
  - Use of stress to signal an important or contrastive word, as in the utterance: 'I wanted BLUE socks' where 'blue' is stressed to suggest that this is an important word, probably contrasting with a previously mentioned colour (Halliday 1967). This use is manifested as differentiation between stressed and unstressed words or syllables by variation in speech rhythm and relative prominence of syllables.
  - Indication of utterance type by the use of intonation, sometimes called inflection (this can also be considered as an interactional or grammatical function) (Cruttenden 1997). For example, a rising intonation at the end of the utterance implies that some response is required, while a falling intonation usually suggests some kind of finality.

- Grammatical functions such as segmenting utterances into phrases by use of pause, stress, intonation and final syllable lengthening (Laver 1994, Cruttenden 1997). For example, in a syntactically ambiguous phrase such as 'chocolate cake and buns' the ambiguity lies in whether 'chocolate cake' is two items or one item (i.e. two simple nouns, 'chocolate' and 'cake' or the compound noun 'chocolate-cake') and this can be differentiated by prosody: primarily, the existence or absence of a pause between the first and second noun, and the duration of the final syllable of the first noun (lengthened if it is last in the 'chunk').
- Affective functions: expression of emotions or of the speaker's affective state by use of intonation and variations in factors such as loudness, speech rate and pitch range that apply to an entire utterance rather than a few syllables of it (Couper-Kuhlen 1986). For example, an utterance said with prosody suggesting positive affect will generally have a wider and higher pitch range than one said with prosody suggesting negative affect.

Prosody in autism

Much of the literature on prosody in autism has focused on prosodic expression for affective or pragmatic purposes and upon the observation that the speech of a child with autism is often characterized by poor inflection and excessive or misassigned stress (Hargrove 1997). It is widely noted by clinicians and researchers alike that the speech of even highly verbal individuals with autism can be 'bizarre' (Fay and Schuler 1980: 31). Descriptions include exaggerated or monotonous intonation, slow syllable-timed speech, a fast rate of speech or an adopted accent different from that of peers (Baron-Cohen and Staunton 1994). Little explanation is offered in the literature for this heterogeneity. Very little is known about the comprehension of prosody in children with autism and what role a receptive deficit might play in prosodic expression. Moreover, prosodic reception is highly relevant to current theories of autism, particularly theory of mind (TOM) (Baron-Cohen 1995) whereby the understanding of a second person's prosody (by an individual with autism) is directly related to his/her ability to infer that person's mental state or pragmatic intent (see 'Affect' below).

To understand the scope of further useful research in this area, it was necessary to review what has already been established with regard to the prosodic ability of children with autism. Although prosody is an area often neglected in therapy for those with autism, disordered prosody can be one of the main barriers to the social acceptance of this population (Shriberg *et al.* 2001). Moreover, prosodic deficits are often life-long even when other areas of language improve.

Methods

A literature search was conducted using the search engines available via First Search, including Medline, CINAHL, PsycINFO, Worldcat, Article First, Proceedings First, MLA, ERIC and Papers First. In addition, LLBA and Web of Science were also searched. Articles published between 1980 and 2002 were considered and the search terms used were (prosody or intonation or accent or stress) and (autism or autistic spectrum disorders or Asperger's syndrome or

pervasive developmental disorder). From this and from references suggested in papers, 16 papers matched the criteria. Two focused primarily on echolalia and one on interpreting affect without reference to prosody. Echolalia was not included as a search term in the main literature search as many papers on echolalia do not deal with prosody.

Results

Table 1 shows the distribution of articles by year of publication and then by topic. The majority of studies were published in the 1980s. The 1990s showed a general paucity of research, but recently there have been a few more studies and Web searches (using the terms 'prosody' and 'autism') show that at least three projects are underway to investigate prosody in autism. Most of the studies covered more than one area of prosody, but stress was the most comprehensively covered area and reception of any type of prosody was seriously under researched.

Table 2 details the number of participants in each study, the type of analysis used and the main aims of each study. One study was a review of the literature on prosodic development in typically developing children and children with autism (Baltaxe and Simmons 1985) and so is not reviewed here. Most studies used very small groups of participants: generalizations from their findings are therefore dubious. Some earlier studies do not adequately define terms such as 'autism' and 'aphasia', making them difficult to compare with more recent studies. Most studies define autism using the DSM criteria but 'Asperger disorder' or 'Asperger's syndrome' were not included in this classification system until DSM-IV (1994). It is therefore possible that some of the studies predating 1994 may have included under the term 'autism' children who would now be classified as having AS, and conversely that children diagnosed with AS before 1994 might in subsequent studies be considered as having HFA.

Table 3 shows the criteria for inclusion in an experimental group in each study. The numbers and types of control participants also varied considerably. Table 4 shows details of the control groups for each study. Types of experimental data varied, including elicited sentences (used for analysis of stress) and spontaneous or elicited conversation. Most studies used perceptual analysis of prosody with only two studies offering instrumental measurements (for details on how to measure prosody acoustically, see Kent and Read 2002: 223–240). Perceptual analysis, of

Table 1.   Distribution by year and topic

Year	No. of studies	Topic	No. of studies
		stress	7
		rate	1
1980–89	7	chunking	5
1990–95	3	affect	2
1996–99	2	reception	5
2000–02	4	echolalia	2
		intonation	2
Total	16		25 (some articles cover more than one topic and so are counted more than once)

Table 2. Participant details for each study

| Study | Year | Domains | Relevant main aims | Participants | | | Controls |
				No.	Diagnosis	Age	
Paccia and Curcio	1982	echolalia	to examine frequency of imitative and contrastive prosody in the echolalia of children with autism	5	autistic-like	6;11–16;10	none
Baltaxe	1984	stress	to examine the differences in use of contrastive stress by children with autism, TD children and children with aphasia	7	autism	4;4–12;2	7 TD and 7 aphasic
Baltaxe et al.	1984	intonation patterns	to explore differences in the intonation contours of children with autism, TD children and aphasic children	5	autism	4;6–12;2	6 TD and 6 aphasic
McCaleb and Prizant	1985	stress	to investigate children with autism's ability to encode new and old information using contrastive stress	4	autism	4;8–14;10	none
Baltaxe and Guthrie	1987	stress	to examine the differences in use of primary sentence stress by children with autism, TD children and children with aphasia	7	autism	4;4–12;3	7 TD and 7 aphasic
Frankel et al.	1987	perception of changes in prosody	to compare the relative intrinsic reward value of differing types of prosody	8	autism	mean 5;7	16 TD and 8 with learning difficulties
Fine et al.	1991	phrasing and stress	to compare patterns of intonation among participants with and without autism	23 AS and 19 HFA	AS and HFA	8–18 and 7–32	34 psychiatric outpatients
Erwin et al.	1991	P3 responses to prosodic stimuli	to determine if people with autism can discriminate prosodic contrasts	11	autism	adults	11 normal adults
Thurber and Tager-Flusberg	1993	phrasing	to investigate the hypothesis that children with autism produce non-grammatical pauses more frequently than other groups	10	autism	mean 12;1	10 TD and 10 with learning difficulties

Table 2. *(Continued).*

Study	Year	Domains	Relevant main aims	Participants			Controls
				No.	Diagnosis	Age	
Local and Wootton	1996	echolalia	a description of the types of echolalia (prosodically echolalic or prosodically contrastive) displayed by a child with autism	1	autism	11	none
Fosnot and Jun	1999	intonation timing and stress	comparing intonation and timing characteristics of children with autism, TD children and children who stutter	4	autism	7–14	4 TD and 4 children with stuttering.
Paul *et al.*	2000	perception	a comparison of receptive prosodic abilities of children with and without autism	18	HFA (males)	not specified	10 TD
Shriberg *et al.*	2001	stress, rate, volume, phrasing and voice	to construct speech and prosody profiles for people with autism and compare them to profiles for people without autism	15 and 15	HFA and AS	10–50	53 TD males
Wang *et al.*	2001	FMRI study of neural correlates of prosody	to determine if people with autism use different cortical networks to process prosody from people without autism	9	autism	mean 9.5	none
Rutherford *et al.*	2002	Comprehension of affect	to determine if people with autism are impaired relative to controls at determining affect from prosody/voice and to relate this to theory of mind	19	HFA or AS	16–59	78 students/staff from a university and 20 other TD adults

Table 3. Criteria for inclusion in studies

Study	Year	Diagnosis	Criteria for diagnosis
Paccia and Curcio	1982	autistic-like	Frequently echolalic, motility disturbances, impaired social relationships
Baltaxe	1984	autism	DSM-III
Baltaxe *et al.*	1984	autism	DSM-III
McCaleb and Prizant	1985	autism	Rutter's (1978) criteria for autism
Baltaxe and Guthrie	1987	autism	DSM-III
Frankel *et al.*	1987	autism	DSM-III
Fine *et al.*	1991	AS and HFA	HFA: DSM-III and full-scale IQ <70 AS: standard assessment (not named)
Erwin *et al.*	1991	autism	not specified
Thurber and Tager-Flusberg	1993	autism	DSM-III-R
Local and Wootton	1996	autism	Childhood Autism Rating Scale (CARS)
Fosnot and Jun	1999	autism	DSM-III
Paul *et al.*	2000	HFA (males)	ADOS and ADI and confirmed by experienced clinicians
Shriberg *et al.*	2001	HFA and AS	DSM-IV for both groups
Wang *et al.*	2001	autism	not specified
Rutherford *et al.*	2002	HFA and AS	ICD-10

course, is useful as it represents easily how 'lay' listeners interpret prosody; but such analysis can be misleading (see in particular 'stress' below).

Discussion

Articles were categorized into the areas of prosody they covered and will be considered in groups according to the prosodic function or topic they addressed.

Stress

'Stress' (here synonymous with 'accent') is described in the literature as 'primary', 'sentential', 'contrastive', etc. Confusion sometimes arises as to what kind of stress is designated by these terms, e.g. primary stress can be contrastive or sentential (Baltaxe and Guthrie 1987). 'Sentential' stress derives from the concept that all complete utterances (spoken 'sentences') contain a stressed item (a word or a syllable), and that if no one part of the utterance is more important than another, the whole utterance is in 'broad focus' (Cruttenden 1997), and the kind of stress it contains is (in this review) termed 'default stress'. Sometimes one part of the utterance is more important (in the speaker's view) than another, and is said to be in 'narrow focus': the kind of stress that is located on this part of the utterance is termed 'contrastive' in this paper; the contrast is usually between a 'new' piece of information and an 'old' or 'given' (previously mentioned) piece (Halliday 1970). Any item in an utterance can be the locus of contrastive stress, including the word/syllable where default stress would occur, which is usually the last lexical item of the utterance (Cruttenden 1997: 75–76 for exceptions). For example, in 'I bought some socks', the utterance is in broad focus (as if in answer to the question 'What did you do?') and therefore contains default stress; whereas 'I bought some

Table 4. Details of control groups

		Controls		
Study	Year	Group One	Group Two	Group Three
Paccia and Curcio	1982			
Baltaxe	1984	7 aphasic matched for MLU	7 TD matched for MLU	
Baltaxe *et al.*	1984	6 aphasic matched for MLU	6 TD matched for MLU	
McCaleb and Prizant	1985			
Baltaxe and Guthrie	1987	7 aphasic matched for MLU	7 TD matched for MLU	
Frankel *et al.*	1987	8 mentally retarded, matched for chronological and mental age	8 TD matched for mental age	8 TD matched for chronological age
Fine *et al.*	1991	34 psychiatric outpatients matched for approximate chronological age		
Erwin *et al.*	1991	11 age-matched normal adults (note: HFA subjects had normal IQ)		
Thurber and Tager-Flusberg	1993	10 mildly retarded, matched for verbal mental age using a receptive vocabulary test and chronological age	10 TD matched for verbal mental age using a receptive vocabulary test	
Local and Wootton	1996			
Fosnot and Jun	1999	4 TD, sex matched and in the same age range	4 children with stuttering, sex matched and in the same age range	
Paul *et al.*	2000	10 TD not specified but most likely chronological age		
Shriberg *et al.*	2001	53 TD adolescents and adults matched for approximate chronological age		
Wang *et al.*	2001			
Rutherford *et al.*	2002	78 adults from the University of Cambridge: not matched	20 typical adults, matched for verbal and non-verbal mental age	

SOCKS' is an utterance with narrow focus, containing contrastive stress, where 'socks' (new information) is emphasized for purposes of contrast, e.g. with a previous mention of 'shoes' (old information), as in answer to the question 'Did you buy shoes?' It is probable that similarity of stress placement in these two utterances is the source of confusion between contrastive and default stress. Contrastive stress is usually manifested by more prominence/salience (greater variation of prosodic forms, e.g. pitch height, pitch movement, loudness) than default stress.

Seven studies have examined the use of default and contrastive stress by children with autism. Early studies in child language acquisition have shown that contrastive stress is established early as a pragmatic device (Hornby and Hass 1970). As such, it is possible that individuals with autism might display deficits in their use of contrastive stress because this population is known to have pragmatic difficulties (Ramberg *et al.* 1996). However, what constitutes 'disordered stress' is difficult to ascertain because the use of stress (or at least contrastive stress) is dependant on the pragmatic intention of the speaker; in a condition such as autism, pragmatic intent is not predictable and difficult to investigate. It is therefore possible that what appears to be a prosodic deficit, in terms of inappropriate or misassigned stress, is in fact a pragmatic deficit. What constitutes a 'stress disorder' has been described as excessive, equal or misplaced stress (Shriberg *et al.* 1990), but the prosodic exponency of these is not well explored in the literature. Misplaced stress would include inappropriate lexical stress (i.e. the wrong use of stress to differentiate class of lexical items, e.g. saying 'IMprint' (noun-type stress) when 'imPRINT' (verb-type stress) was needed), as well as emphasis of the 'wrong' word in an utterance. Excessive stress would include prominence on inappropriate words (e.g. grammatical function words) or on syllables that would normally be unstressed as well as on appropriate ones, although it should be noted that multiple stresses in an utterance are normal in the emphatic speech of the unimpaired population, but in this case the stresses occur on 'stressable' syllables. Lack of stress is harder to define: it may denote minimal prosodic variation in whole utterances, as well as on stressed items; but again, this is not uncommon in unimpaired speech, and normal perception of language tends to discern some form of stress even where there is very little to show for it in terms of prosodic exponents.

Disordered receptive abilities regarding stress may be more straightforward to identify. A listener seldom needs to be aware of default stress, but needs to be able to identify the place of contrastive stress in order to interpret the pragmatic content of the speaker's utterance. This has implications for TOM in that the listener must infer why the speaker stressed a particular word or words in order to interpret the utterance.

McCaleb and Prizant (1985)

McCaleb and Prizant investigated the use of contrastive stress in four children with autism. The children were males aged 4;8–14;10 and met Rutter's (1978) criteria for autism, which are similar to the DSM-III criteria (1980). Mean length of utterance (MLU) measures for the children ranged from 1.96 to 2.82, but no information was available about non-verbal ability or receptive language ability. Contrastive stress in this context was the marking (stressing) of new (versus old) information in spontaneous conversation. Several studies of stress acquisition (Hornby and Hass 1970,

MacWhinney and Bates 1978, Fletcher 1985) conclude that typically developing children use contrastive stress appropriately for new information by the age of 3.

The data consisted of two 20–30 minutes of videotaped interaction with the child's speech and language therapist and with his/her class teacher (Teacher: 'Yes, you found your own toothbrush'. Subject 2: 'TOOTHbrush. Brian TOOTHbrush'; compare with 'Man. BLUE man' in Wieman's (1976) study of typically developing children). All referential utterances were then coded as either new or old information and then analysed for lexicalization and contrastive stress in line with Greenfield's (1979) criteria. Results indicate that all of the participants stressed new information as often as they stressed old information and that they used lexicalization and contrastive stress equally often.

The authors conclude that the children were using stress in an atypical way. Transcriptions of examples of the data suggest, however, the possibility that in 'stressing' old information the children could have been using stress correctly according to what they wanted it to do, e.g. (re)stress old information because to them it still had relevance to the exchange. If they were failing to follow the development of a conversation, persisting instead with their own conversational priorities, this would be consistent with the pragmatic deficit often reported in children with autism; but the encoding of old information would not be incorrect in this circumstance, although the children's stressing may have been atypical in degree or manner. No comprehension measures were taken and it is therefore not possible to know if the prosodic output problem (if there is one) was due to an underlying receptive deficit.

Baltaxe (1984)

Baltaxe investigated the use of contrastive stress in seven children with autism, seven aphasic (severe language delay or disorder and normal non-verbal IQ) and seven typically developing (TD) children aged 2;9–12;2. The children with autism conformed to the DSM-III (1980) criteria but no information was given regarding their non-verbal abilities or general language levels. The groups were matched by expressive language level as determined by MLU, which ranged from 1.9 to 4.9 morphemes.

Data were obtained using yes/no questions that were counterfactual to a play situation. Contrastive stress was elicited for subject, verb and object positions for all of the stimuli. For example, to elicit contrastive stress in the subject position, the participant was shown a doll named Pat sitting on a chair and asked 'Is Mike sitting on the chair?' The expected response is therefore 'No, *Pat* is sitting on the chair' (with focus or stress on 'Pat'). Two skilled listeners perceptually analysed these data for incidences of contrastive stress. Responses were deemed to contain contrastive stress when the testee responded using a subject–verb–object (SVO) phrase rather than the more natural 'yes' or 'no'. All groups provided responses only 60% of the time (presumably the remainder of responses were simply yes/no replies). Of these, the typically developing children displayed the highest percentage of prosodically correct utterances followed by the aphasic children and then the children with autism. Children with autism were perceived as misassigning stress twice as often as the typically developing children. Most notable was their tendency to assign stress to more than one stressable syllable (36.8% of all their errors), an error type which the other children did not produce at all.

Too little information about data and judgement tasks is given to be sure what constituted prosodic correctness in this experiment, but it seems fair to suppose that much depended on the children's ability to understand the requirements of the task and to cooperate with their artificiality, and that not surprisingly children without a pragmatic deficit coped better. It would be unwise to conclude that the findings of this experiment reflect the natural or conversational use of stress in children with autism.

Baltaxe and Guthrie (1987)

In a sister study to Baltaxe (1984), Baltaxe and Guthrie investigated the ability of the same children to use default stress. An elicited production task in which the examiner manipulated toys while asking 'What's happening?' was used to obtain SVO and subject–verb–preposition (SVPrep) sentences with default stress. Sixteen utterances consisting of eight nouns and four verbs were recorded. A correct response was an utterance with stress on the last stressable syllable. For example, in the sentence 'the baby is sleeping on the bed', one would expect 'bed' to be stressed. The pattern of scoring was similar to that found in Baltaxe's contrastive stress study: the typically developing children scored highest followed by the aphasic children and then the children with autism, again suggesting to the authors that the language impaired (aphasic) children and the children with autism had some difficulty with this type of prosody. This cannot be explained by differing language levels as all the children were said to have similar language abilities as determined by mean length of utterance measures; nor can it be explained by the artificiality of the task.

However, all groups made more incorrect than correct responses, with even the typically developing children assigning stress correctly in only 20% of their responses, and the children with autism made no correct responses. One hypothesis to account for this was that default stress may not be reliably used in even adults' speech. This was tested by asking eight adults to read the sentences using a neutral intonation. Results indicated that the adults, in fact, were reliably using default stress suggesting that the children had not yet reached the adult form. Nevertheless, there is a methodological problem in that differences between prosody for reading and for conversation (spontaneous and elicited) are not taken into account. In particular, Howell and Karima (1991), in a comparative study of prosody for reading and spontaneous speech, note that in their adult participants the position of stress differed between reading and spontaneous conditions.

Another hypothesis is that default stress is not yet established in typically developing children of less than 4 years; but the studies of typically developing children already cited belie this, suggesting that default stress is established before contrastive stress: children younger than 3 years show a preference for locating stress at the end of an utterance.

In error patterns for the children, the authors found that stress was misassigned to the first stressable element of the sentence (the subject noun) approximately 87% of the time for all groups. They propose that stressing the subject may reflect the fact that the children were at the egocentric stage of development and that to them the subject (agent) is the most important element in the phrase, that is, the children are underlining its pragmatic function. On the other hand, this inter-pretation contradicts findings in Baltaxe (1984) that erroneous stress placement

tended to involve a shift rightwards from the target, away from subject position, to a subsequent stressable element of the phrase; as well as results from other experimental and observational studies (Wells and Local 1993 for a review), which suggest that where children make errors with stress placement, it is by shifting it to the last word in the utterance (e.g. Hargrove and Sheran 1989).

The finding that typically developing children did not place default stress correctly may be an artefact of the scoring of the task, but too little information is given about this process.

Fine et al. *(1991)*

Fine *et al.* investigated intonation in 23 individuals with AS (aged 8–18), 19 with HFA (aged 7–32) and 34 psychiatric outpatients (aged 7–18, acting as controls). Data consisted of a 10-minute sample of speech, considered to be representative of spontaneous conversation, in which an investigator interviewed the participant on topics such as family and school. This was then coded by a professional phonetician who perceptually marked the tone boundaries and the stress assignment, which were then in turn judged as either appropriate or inappropriate by a research assistant.

The authors concluded that those with HFA were less communicatively effective in their use of intonation than the other two groups, and that this may be an effect of the deviant language development that is typical of children with autism but not of children with AS. The authors report that the participants with AS performed as well as the controls and that therefore disordered intonation is a manifestation of disordered language.

Nevertheless, the participants with autism showed appropriate default stress on the last content word of the tone group (a finding at odds with Baltaxe and Guthrie 1987). Contrastive stress was reportedly less often well placed; the participants with autism instead tended to place stress on function words rather than on content words, thus employing stress in a way that was not communicatively useful.

Fosnot and Jun (1999)

Fosnot and Jun investigated the prosody of four children with autism, four children who stutter and four typically developing controls. Participants were boys aged 7–14 years with no matching criteria specified. The children were required to read eight sentences three times and then, on a separate occasion, imitate the sentences as produced by an adult. The sentences were as follows (Fosnot and Jun 1999: 1926):

   (1) It's a rhino.
   (2) It's a rhino?
   (3) It's not a rhino.
   (4) It's not a rhino?
   (5) It's a rhinoceros.
   (6) It's a rhinoceros?
   (7) It's not a rhinoceros.
   (8) It's not a rhinoceros?

Regarding results relevant to stress, the authors found that the children with autism produced more pitch accents (main stresses) than the other two groups of children. Word stress and sentence stress were both misassigned, for example, in the word 'rhino' the first syllable is usually stressed but the authors report instances of the second syllable being stressed and some function words were also inappropriately stressed. This study has the same disadvantage as Baltaxe and Guthrie (1987) in that it examines only reading and imitative prosody. Results were analysed acoustically for pitch range and duration and perceptually for types of pitch accents and the location and type of boundary tones, and it is not reported how well the instrumental measurements concurred with the perceptual judgements, or which type was preferred in the case of discrepancy.

Shriberg et al. *(2001)*

In a more recent study, Shriberg *et al.* investigated the speech and prosody characteristics of 15 male speakers with HFA, 15 male speakers with AS and 53 typically developing males. All participants were in the wide age range 10–50 years. The HFA and AS groups both conformed to the DSM-IV (1994) criteria for their respective diagnoses and information about non-verbal ability and expressive and receptive language abilities was included.

The authors used the 'Prosody-Voice Screening Profile' (PVSP; Shriberg *et al.* 1990), an assessment based on output skills only, to rate a conversation sample. For the AS and HFA participants, this conversation sample was a videotaped standardized clinical interview and for the controls the content of the conversation sample was not specified. Stress was only one of the domains covered with inappropriate lexical stress and inappropriate phrasal (default) stress being considered. Speakers in the control group displayed appropriate stress (both lexical and phrasal) in 95.2% of utterances; AS speakers in 86.5% and HFA participants in 77.3%. Therefore, although the HFA participants used stress appropriately in the majority of utterances there was still some evidence of difficulty with contrastive stress. However, these results are difficult to generalize due to the wide age range of participants: no analysis was offered for these age differences, and no measure of language-age was taken, so it is possible that some of the differences reflect different stages of prosodic development.

Paul et al. *(2000)*

It can be seen that much of the literature has covered prosodic production making it difficult to know if expressive deficits have underlying receptive prosodic deficits. However, in a small pilot study (note that only limited information was available for this study), Paul *et al.* describe an investigation of the perception (although comprehension and not simply acoustic perception was involved) of grammatical and pragmatic aspects of prosody in 18 males (age not specified, but probably children) with HFA compared with 10 controls aged 12–18. The stress perception task for both the grammatical and pragmatic aspects (it is not made clear how these tests were differentiated) required participants to listen to single words differentiated by stress and make judgements about the class of that word: the noun/verb stress distinction, as in 'IMprint' the noun versus 'imPRINT' the verb).

The selection criteria for participants are not available and the small number of

controls make it difficult to draw conclusions. Moreover, no information is available about the prosodic production abilities of the children. Results showed that the participants with HFA were less able to comprehend this type of stress, although no information about statistical significance is available and the difference between pragmatic stress and grammatical stress is not made clear. Results are further complicated as the authors did not check for understanding of the words used.

Summary of findings on stress production in autism

Individuals with autism do appear to express stress atypically, with contrastive stress judged problematic in all studies that investigated it. Findings about default stress conflict, with two studies finding problems and two finding none.

The nature of atypical stress is unclear, however, first in terms of processing level. If speakers perceptibly stress what they want to stress but the resulting stress patterns fail to concur with expectations, this is not a prosodic deficit, or stress misassignment, but a manifestation in speech of a different problem (e.g. pragmatic deficit). At the functional level, a 'wrong rule' may be being applied, i.e. a speaker might have inferred a prosodic stress system at odds with the linguistic environment (suggesting perhaps a receptive deficit); but to ascertain expressive prosodic deficit reliably, an experiment must establish that speakers are failing to assign stress where they intended, or assigning it to items where they did not intend it; in the studies reviewed here, this may or may not be the case.

Conditions under which data were obtained, particularly contextual, are seldom satisfactorily fleshed out, in these studies. Where they are, it appears that there might often be reasons for children (especially those with autism, who may have atypical dialogue agendas) to wish to change stress placement from where the experimenters expected it to occur. If intention is impossible to ascertain (as is likely in conversational samples), performance should be compared with that of a statistically valid number of typically developing peers; and the utterances involved, as well as the target stress patterns, should be ecologically valid (e.g. not expected to include a predicate that can be pronominalized, as in Baltaxe 1984).

However, it might be the case that speakers with autism assign stress unintentionally, i.e. have a problem at the execution level. To evaluate this, an experiment should also examine (instrumentally) the prosodic exponents not only of the stressed item, but also of the rest of the utterance, and establish (by reference to typically developing, sociolinguistic and language-age peers) the coordinates of prosodic exponency and stress for the neurotypical population and in what respects children with autism diverge from these. If this could be established, it would be an advance in knowledge of language acquisition and an advantage for speech and language therapy.

A problem with perceptual evaluation is that it is not clear in any of the studies that used it whether judges were aware of or had heard the stimuli that preceded the responses they assessed for stress assignment. If they knew the stimuli, they would have expectations as to what stress pattern would be appropriate (Couper-Kuhlen 1986), and this might influence their perceptions. Admittedly, the effect would probably be to incline perceptions towards appropriacy rather than away from it, but it would still say more about the judges' perceptions than about the speaker's stress patterns. This problem would not be resolved by inter-rater reliability tests, since all judges would be subject to the same expectations. It is

asserted in some of the studies that the transcribers and judges were linguistically or phonetically trained, and might therefore be able to divorce expectation from judgement, but this in turn would not answer the question of whether the stress patterns would be perceived as inappropriate in everyday social contexts.

Fourth, there is the question of how valid, currently, are the prosodic tenets of information structure: that new material will be stressed and old material 'de-stressed', or that default stress will be manifested by declination with non-prominent pitch movement on the final item. Diachronically, the deployment of stress is subject to change; and if, for example, the exponency of final contrastive stress is frequently indistinguishable from the exponency of default stress (Peppé *et al.* 2000), either the principles of information structure or the exponency of stress may be in the process of changing. If so, the new exponency is what children will learn, and methods of assessing prosodic deficit will have to take this into account.

Finally, many of the studies comparing typically developing and language-impaired (or 'aphasic') children with autism use inadequate language measures and therefore it is difficult to know if prosodic difficulties represent a more general language deficit.

Phrasing and chunking

Fewer studies have investigated verbal phrasing or boundary prosody, here designated 'chunking' to distinguish it from syntactic or written phrasing. Chunking is the prosodic segmentation of utterances for grammatical, pragmatic or semantic purposes. This is most easily illustrated by syntactically ambiguous sentences such as the example given above, in which the ambiguity is resolved by duration and intonation cues. Although the majority of utterances do not rely heavily on prosodic cues for intelligibility, and indeed syntactically ambiguous sentences are often disambiguated by context, prosodic boundary cues are an important aid to decoding (a good illustration of this is to imagine trying to read a page of text with no punctuation), and for making decisions about whether or not a speaker has finished his/her turn. Receptively, a deficit in chunking would perhaps manifest itself as a lack of awareness that a speaker is finished and an inability to understand syntactic groupings in speech where distinguished by prosody alone. Expressively, an individual may interrupt speakers, or not use prosodic boundaries to indicate coherent chunks (Shriberg *et al.* 2001), thus leading to confusion on the part of the listener.

'Phrasing' is sometimes used interchangeably with chunking, but 'phrasing' can be interpreted as including semantic features. Shriberg *et al.* (1990) define phrasing as a flow of word and phrase groups, which apparently concerns prosody only; but in their 'Prosody-Voice Screening Profile', inappropriate phrasing is then defined as repetitions, revisions and inappropriate rate, all factors that concern lexical content and dysfluency issues as well as the completeness or otherwise of prosodic chunks.

Fine et al. *(1991)*

In the study by Fine *et al.*, the participants with autism were able to place boundary tones at the end of intonational phrases successfully, suggesting an ability to phrase prosodically (see above for more details regarding methodology and analysis). As for the investigation of stress, there was no mechanism in this study for determining

how far the impression of chunking matched speaker intention. It was found, however, that non-grammatical (e.g. hesitational) pauses were used less frequently.

Thurber and Tager-Flusberg (1993)

Thurber and Tager-Flusberg investigated pauses in narrative speech by comparing grammatical (pauses at phrase boundaries) and non-grammatical pauses (pauses within phrases) in 10 children with autism (mean age 12;1), 10 children with learning difficulties (mean age 11;3) and 10 typically developing children (mean age 7;9). The children with learning difficulties were described as 'mildly retarded', but were educated in specialist provision, and the children with autism were diagnosed in line with the DSM-III-R (1987) criteria. The purpose of the experiment was to investigate communicative or cognitive load, which is reflected by frequency of pauses. Previous studies have reported that typical adults exhibit a greater frequency of non-grammatical pauses in cognitively demanding tasks compared with less demanding tasks (Goldman-Eisler 1972). Non-grammatical pauses are therefore an indication of cognitive processing and are reflective of such processes as lexical decision.

A narrative speech sample was elicited using a wordless picture book. This was then coded for silent pauses, filled pauses, false starts and repeats; with a distinction being made between grammatical and non-grammatical pauses. Results confirm those of Fine *et al.* (1991): the children with autism made use of grammatical pausing in a similar way to the typically developing children and the children with learning difficulties. Additionally, Thurber and Tager-Flusberg found that the children with autism used non-grammatical pauses less frequently than the other groups. The authors suggest that this finding is an indication of reduced cognitive load and attribute it to less communicative investment in the interaction.

Fosnot and Jun (1999)

In direct contrast to Thurber and Tager-Flusberg (1993), Fosnot and Jun report that the children with autism in their study were more likely than typically developing children to put in non-grammatical pauses, e.g. after 'It's' in the sentence: 'It's a rhino', in the reading condition. There is the problem of non-matching experimental method (reading versus non-reading prosody), but Fosnot and Jun's read sentences should contain even less cognitive load than the narrative speech sample collected by Thurber and Tager-Flusberg. Fosnot and Jun found a similar pattern in the imitation data and they conclude that the children with autism were unable to imitate the timing and chunking patterns of adults. However, this is problematic, in that it does not appear that the children were specifically asked to imitate the stimuli: they were asked to repeat it, but not explicitly to imitate the prosodic properties. Moreover, some studies of echolalia (e.g. Local and Wootton 1996) report children with autism spontaneously echoing the prosody of utterances, suggesting that this is possible for at least some children with autism.

Shriberg et al. (2001)

These authors reported that 40% of their speakers with HFA had inappropriate or dysfluent phrasing on more than 20% of their utterances, thus apparently

contradicting the findings of Fine *et al.* (1991) and of Thurber and Tager-Flusberg (1993). Closer inspection reveals, however, that the error types were sound, syllable and word repetitions, i.e. indicative of content dysfluency, arising perhaps from word-finding difficulty. For these problems, incomplete chunking would be appropriate. This study is therefore a good example of the need to distinguish clearly between prosody and other linguistic parameters in speech, and the confusion that can arise from not doing so. With this distinction in place, Shriberg *et al.*'s findings can be seen not as contrary to the other studies but as supporting an implication of them: that children with autism not only chunk effectively when fluent, but also that they display incomplete chunking when appropriate, i.e. when dysfluent.

Paul et al. *(2000)*

This study assessed the comprehension of grammatical phrasing. A chunking task was used in which listeners were required to appreciate durational and intonational cues at phrase boundaries, e.g. in the difference between 'Paul, my friend, is here' and 'Paul, my friend is here'. Results show that in contrast to the studies assessing expression of phrasing, the participants with HFA did less well in this task than the controls. However, no statistical information is available to determine how significant the difference was, and no assessment of expressive ability was carried out.

Summary of findings on phrasing in autism

At first sight, the overall picture of prosodic phrasing in autism, minimal though it is, is contradictory: it appears that the participants in Fine *et al.*'s (1991) and Thurber and Tager-Flusberg's (1993) studies displayed typical prosodic phrasing, while those in Fosnot and Jun's (1999) and Shriberg *et al.*'s (2001) studies did not, although the latter may have displayed prosody appropriate to dysfluency.

The impression even from so few studies and participants is likely to be unreliable for several reasons. Only one aspect of prosodic phrasing (pauses) is considered, and not the other relatively well-established exponents: the duration as well as the occurrence of pauses (Kowal *et al.* 1983), final syllable lengthening (Cruttenden 1997) and F0 change (pitch movement) (Schuetze-Coburn *et al.* 1991). The place of pauses within utterances is not considered, but this has a bearing on whether a pause is perceived as a 'legitimate' space for lexical retrieval (e.g. between determiners and lexical items), and thus characteristic of typical speech and prosody, or as an inexplicable pause (e.g. mid-word), suggesting atypicality. Terminological differences are apparent: for Shriberg *et al.* a phrase is seen as what occurs between pauses in speech, while for Thurber and Tager-Flusberg pauses can be seen as interrupting a phrase. Except in Fosnot and Jun's study, the data consisted of conversation samples where there is no control for how much judges were 'expecting' correct phrasing and therefore heard it; notably, the single comprehension study used task-based assessment, involving minimal pairs of utterances. No measurements of the exponency of prosodic phrasing are reported to support the listeners' perceptions.

Affect

Two studies have specifically addressed the issue of comprehension and expression of prosodic affect in autism. However, in other studies, the communication of affect has been investigated within domains other than prosody, e.g. facial expression, gesture, voice and face recognition (Hobson 1986a, b, Boucher *et al.* 1998). Often, these studies are conducted from a psychological rather than from a linguistic point of view. They therefore give interesting results about the use of prosody pragmatically but they often do not explore prosody as an aspect of language or do not control stimuli linguistically. Findings from these studies confirm that individuals with autism have difficulty interpreting affect, and Boucher *et al.* (2000) report an inability to name vocally expressed affect. Studies in which many non-prosodic factors are involved will not be described in detail as it is difficult to draw conclusions about the role of prosody in these experiments.

Nevertheless, prosody plays an important role in the understanding and expression of affect. Positive affect is generally associated with higher and wider pitch range and altered rate, and negative affect with lower and narrow pitch range (Kent and Read 2002). Difficulties with social or emotional reciprocity are a defining feature of autism (DSM-IV 1994), and a difficulty with understanding or expressing affective prosody may be part of this. A receptive deficit could be diagnosed if an individual was unable to interpret emotion based on prosody alone; and an expressive deficit could be diagnosed if there was an inability to convey affect using only prosody.

Rutherford et al. (2002)

Rutherford *et al.* investigated the ability of 19 adults with HFA or AS to judge the affective meaning of 40 phrases; the authors view this experiment in the context of TOM. The stimuli consisted of segments of dialogue recorded from dramatic audio books, varying in more than just prosody (e.g. in vocal quality, articulatory settings and such 'paralinguistic' features as loudness, speech rate and pitch range). Participants were required to listen to stimuli and then decide which of two adjectives best described the affective content. An example is the stimulus 'Keep the damn thing' with the choice of 'irritated' or 'surprised' as adjectives to describe the utterance. Results showed that the HFA and AS group was impaired on this task compared with the performance of a large number of typical adults, and that the impairment did not correlate with verbal or performance IQs.

Methodological drawbacks to this experiment leave its conclusions open to question. There are some relatively low scores among the typical participants and these results are not accounted for. Some of the adjectives were low frequency (e.g. 'derogatory', 'accusatory', 'intrigued', 'contemplative') and it is possible that the understanding of these terms was variable; the only check on this was that the participants (who all had normal intelligence) were given the opportunity to request clarification, and that none did. Some of the stimuli may have been semantically biased towards one of the answers, for example: 'I'm afraid he's gone out, sir' semantically favours 'apologetic' rather than 'hurried'. This is another example where communication effects—or lack of them—are attributed to prosody (and, in this case, voice features) alone, when some should be attributed to lexis or other linguistic parameters.

Paul et al. (2000)

Receptive affective prosody is also addressed by Paul *et al*. In what the authors term 'pragmatic phrasing', the understanding of the difference between utterances said in an 'excited' or 'calm' register was assessed. Both the children with HFA and the typically developing children performed near the ceiling in this assessment. This may suggest that children with HFA have no problem understanding affective prosody (contrary to the findings of Rutherford *et al*. 2002), but may also be because the difference between these two emotions is substantial and a difference might have been found between groups if the stimuli had been emotions that were more complex or subtle.

Intonation patterns

Whereas the communication areas we have considered thus far might be thought of as communicative functions (phrasing, stress, affect), 'intonation patterns' rather denote the means by which such functions may be achieved. Intonation was included as a search term because some studies approach the topic by examining the means rather than the communicative effect.

Although intonation plays a subtle role in many aspects of affect as well as interaction, the most specific use of intonation found in our search deals with the fairly crude distinction that questions are produced with rises and statements with falls. In practice, the pitch height and range of the utterance, as well as pitch movements such as rise or fall, are distinguishing features of utterance type. For example, utterances ending with a rise starting at low pitch and covering a narrow range are likely to be heard as indications of more to follow (from the speaker): 'cake' said with a low narrow rise could be one of a number of foods being listed. For a rising utterance to be heard as a question, i.e. requesting a response from the interlocutor, the rise needs to start high and rise steeply: 'cake?' (Cruttenden 1997, Peppé 1998). With an inventory merely of rises and falls, it is therefore hard to quantify the 'disorder' of utterances that are deemed to be intonationally disordered.

Baltaxe et al. (1984)

Baltaxe *et al*. analysed the intonation contours of five children with autism (aged 4;6–12;2), six children with aphasia (4;5–12;2) and six typically developing children (2;0–4;0); MLU was 1.45–4.46 for all groups. Spontaneous declarative utterances (SVO) were recorded under controlled conditions and analysed (acoustically) for frequency range, terminal fall, intonation contour, declination effect, and covariance of frequency and intensity.

Results were based on acoustic measurements and show that for frequency range the typically developing children had the greatest range, then the children with autism and then the aphasic children. This difference was only significant between the TD and the aphasic children, so the children with autism did not have a significantly different frequency range from the typically developing children. However, the children with autism presented with either very narrow or very wide ranges. A similar pattern was reported for all the other parameters. This suggests that the mean of the frequency ranges did not adequately capture the atypicality of

the children with autism. Apart from methodological reservations (the findings were derived from only six children, verbal mental age was not apparently taken into account), it appeared in general that children with autism displayed less stability and greater individual variation than the typically developing peers.

Fosnot and Jun (1999)

The stimuli used by Fosnot and Jun consist of minimal pairs of interrogative and declarative utterances, not differentiated by lexis or syntactic form, with the expectation that the declaratives would be produced with a fall or rise–fall at the end, interrogatives with a rise or fall–rise. In this study, the authors were expecting children to take note of the question mark in the written stimuli and produce prosodic minimal pairs accordingly. However, they found that the children with autism, unlike the other two groups, did not make a distinction between the interrogatives and declaratives, in that both sounded like declaratives. The children were all described as 'sight readers', but this does not explore whether or not the children were actually aware of the function of the '?' in the written sentences.

   The performance of the children with autism was better in the imitation condition, showing that they at least have the ability to produce this type of intonation pattern, although they may chose not to use it or understand its communicative function. Interestingly, the authors found that the ability to imitate this type of intonation pattern correlated with severity of autism and length of sentence. The authors therefore claim that the ability of children with autism to produce prosodically correct features is a measure of severity of autism. This is a promising finding for the relevance of testing prosodic ability in children with autism, but as it is only based on four subjects and the authors do not detail how they measured severity of autism it should be interpreted with caution and perhaps thought of as an area for future research.

Perception of changes in prosody

Infants are highly sensitive to changes in ambient prosody; at only 4 days old, a child can distinguish a familiar language from an unfamiliar language based on prosody alone (Price *et al.* 1991). The ability to perceive changes in prosody is obviously dependent on auditory discrimination skills, which are therefore neces-sary if one is to understand prosody. A deficit in perceiving changes in prosody would therefore have far reaching implications for receptive prosodic ability, and possibly a consequent effect on production ability. In experiments investigating this, however, auditory memory skills must also be considered.

Frankel et al. (1987)

Frankel *et al.* used an operant reinforcement paradigm to investigate the reward value of prosodic features for different groups of children. Participants were four groups of eight children: one with autism (in line with DSM-III criteria); one with learning difficulties ('mentally retarded'), matched to the group with autism for chronological age and mental age; and two groups of typically developing children,

one matched to the two experimental groups for mental age and the other for chronological age.

Stimuli were four prosodically different versions of the children's story 'The Three Bears'. The first condition was 'natural' in that it had prosody that one would typically associate with reading a story to a young child. The next condition was 'monotone', with an attempt being made by the speaker to maintain a constant fundamental frequency (note the stimuli were not resynthesized and therefore not truly monotone); the third condition was 'staccato', with all syllables stressed with a staccato rhythm. In the last condition, 'metronome', the story was prosodically monotone and unstressed.

The participants were conditioned to pull a lever in a small experiment room. Pulling the lever resulted in 3 seconds of the story being played, therefore more frequent lever pulling was thought to suggest greater interest in the story or in the prosody with which the story was being told. Each prosodic condition of the story was presented for five 20-second periods, with each new prosodic condition resulting in continuation of the story.

Results indicate that none of the groups showed statistically significant differences in frequency of lever pulling for any of the prosodic conditions. The authors suggest that this is evidence that the children with autism did not have a prosodic input preference, but of course neither did the typically developing children. The typically developing children showed high rates of lever pulling for all prosodic conditions, perhaps suggesting that they were interested in the content (rather than the prosody itself) of the story; or that the stimuli had a novel value. However, the children with autism and the children with learning difficulties only demonstrated high rates of lever pulling during the first 20–40 seconds of each new prosodic condition. This, then, suggests that the children with autism were at least aware of the change in prosody. Of course, this study tells us nothing of the ways in which children with autism understand prosody.

Neural processing of prosody

Much of the research in prosodic ability in autism focuses on the behavioural or observable aspects of prosody with little or no reference to the neural processes, or differences in the brains of individuals with autism, that may underlie prosodic deficits. It is generally thought that affective prosody is literalized to the right hemisphere and linguistic prosody to the left (Baum and Pell 1999). The notion of linguistic and affective prosody as discrete entities is controversial (Seddoh 2002), but given that there is some evidence that the basis for autism may be attributed to right hemisphere dysfunction (Sheilds *et al.* 1996), this would perhaps suggest that one would expect individuals with autism to have more difficulty with affective prosody than with linguistic prosody. However, the neural basis of prosody in typical adults remains controversial (Baum and Pell 1999), and most of the evidence comes from studies using adults with localized lesions as participants, making it difficult to draw conclusions about the neural processing of prosody in children with developmental disorders. Nevertheless, with the advent of neuroimaging techniques such as functional magnetic resonance imaging (fMRI), it is possible to investigate neural processing of prosody in individuals with autism.

Erwin et al. *(1991)*

Erwin *et al.* investigated the P3 responses (a P3 response is a positive wave on an electroencephalogram (EEG) peaking at around 300 ms after task-relevant stimuli) to prosodic stimuli for 11 adults with HFA and 11 age-matched controls.

The cognitive P3 potential was recorded in response to three types of stimuli: phonemic ('ba' versus 'pa'), linguistic-prosodic ('Bob' said as a statement versus 'Bob?' said as a question) and emotional-prosodic ('Bob' said with 'happy' prosody versus 'Bob' said with 'angry' prosody). Auditory discrimination was also considered by asking participants to press a button in response to certain stimuli and by asking them to match the stimuli to an appropriate picture.

Results showed that all participants displayed normal P3 responses to all of the stimuli and performed within normal limits for the auditory discrimination component of the task. This suggests that, at least at the biological level, adults with autism are processing prosody in a way similar to typical adults, although findings might be unreliable as the group of participants was small. If adults with autism process prosody normally, this perhaps gives weight to the argument that children with prosodic problems may be presenting with delayed rather than disordered prosodic patterns; such problems are nevertheless usually persistent. Given that disordered expressive prosody is not a constant in children (or adults) with autism, its presence or otherwise might be related to differences of neural processing, and it would have been interesting to know whether adults in this study were affected in this way.

Wang et al. *(2001)*

Wang *et al.* used fMRI to investigate the processing of affective and linguistic prosody in nine boys (mean age 9;5) with autism spectrum disorder. The authors used a same/different paradigm where the participants were required to listen to pairs of sentences and decide whether they sounded the same or different. Eight pairs of sentences were constructed, half with 'neutral intonation' and half with differing prosodic cues. For linguistic prosody, half had a rising intonation at the end of the sentence (as is typical of a question); for affective prosody, half the stimuli had prosody suggesting sad affect and half-angry affect. In addition, a semantic control condition was constructed in which participants had to decide if sentences had the same meaning.

For each participant, 84 functional images were acquired. Results show that all the participants performed above chance. For the affective prosody condition, selective activation in the middle and superior temporal gyri in the right hemisphere was observed (consistent with the idea that affective prosody is processed in the right hemisphere). For the linguistic prosody condition, a few small clusters of activity were observed in the left middle temporal and inferior frontal gyri. According to the authors, the participants demonstrated a pattern of lateralization similar to typical individuals, but the specific regions of activation differed. The children with autism in this study might therefore have been processing prosody in a different way to typical children. This could account for difficulties with understanding prosody, perhaps leading to difficulty using expressive prosody.

Echolalia

Echolalia is the repetition (echoing) of utterances made by another person (Prizant and Duchan 1981). Although most studies of echolalia are not directly relevant to prosody, one study that investigated echolalia in autism specifically addressed the issue of whether or not echoed responses are prosodically matched. Local and Wootton (1996) present a single case study of an 11-year-old boy, 'Kevin'. Kevin is described as having 'severe autism' as quantified by The Childhood Autism Rating Scale (CARS; Schopler *et al.* 1986). This in itself is interesting as most studies of prosodic ability in autism focus on high-functioning individuals. The investigation focuses on Kevin's use of immediate echolalia to communicate with, in the main, his mother. Local and Wootton explore the communicative value and the prosodic characteristics of his echolalia.

Kevin does not generally initiate conversation and when asked questions he does not give verbal responses immediately. The authors describe as 'unusual echoing' echoes that are prosodically and segmentally echoed and occur immediately after the model. This type of echo is not present in typical development. Paccia and Curcio (1982) suggest that individuals with autism are more likely to use such echoes after questions that are beyond their understanding; Local and Wootton (1996), however, argue that there is little evidence that the questions are ungraspable by Kevin, so they find his use of echo functionally opaque. The authors go on to suggest that Kevin is choosing repetition as a response because it is often a successful conversational tool for him.

It is, however, also possible that Kevin's echoing, far from being functionally opaque, is a normal use of prosodic ability to help interpret what he hears. When the whole of an incoming utterance is unintelligible, it would be appropriate for Kevin (or anybody) to echo back exactly what has been heard rather than manipulate prosody to stress a single item, and it is precisely such a lack of stress variation that would be the evidence that the question was ungraspable, as suggested by Paccia and Curcio (1982). Kevin is not incapable of sometimes stressing single previously unmentioned items, suggesting that appropriate manipulation of prosody is available to him. For a reason why Local and Wootton (1996) found his echoing unusual, it therefore seems advisable to look further than prosody.

Summary and conclusions

There are contradictions in the findings in all the areas of prosodic function identified in this review. The big questions are whether these contradictions are the result of discrepancy of subject (i.e. some subgroups of ASD may present with one type of prosodic disorder, some with another) or the result of discrepancy of investigative method, but so much of the latter has been found that it is impossible to reach any conclusion about the former.

No study offers a large number of subjects, matched with typically developing children or adults (controlled for linguistic and non-verbal abilities). If findings were consistent, small-scale studies would offer pointers, but as it is these do not inspire confidence. Moreover, those studies that include more than 20 individuals with autism or AS (Fine *et al.* 1991, Shriberg *et al.* 2001) present groups with wide age ranges and more than one diagnosis. Only two studies (Baltaxe *et al.* 1984, Fosnot and Jun 1999) use acoustic analysis to quantify expressive prosody (seven

do not); more is needed to establish the prosodic features that characterize both atypical and typical prosody. Uncertainty of definition in some areas of prosody (such as stress) appeared to be responsible for some conflicting findings, as did the attribution of some disorders to prosody when other communication parameters were implicated. Only one study (Shriberg *et al.* 2001) used a published and readily available prosody assessment, and no studies use a standardized prosody assessment, i.e. one that has been used to gather a large amount of normative data with which to compare data from people with disordered prosody. In several areas of prosody (such as affective and interactional prosody), there are very few studies. In particular, the understanding of prosody is seriously under researched: very few studies address this, and none in the context of expressive abilities. Furthermore, the fact that no study covers a wide range of both expressive and receptive prosodic abilities makes it impossible to investigate issues concerning the relationship between receptive and expressive ability. This is particularly relevant to current research in cognitive theories of autism, particularly TOM. It is possible that an understanding of the difficulties that people with autism have with interpreting incoming prosody (particularly, but not only, affective prosody) might lead to greater understanding of why these individuals have TOM problems. Add to this the fact that unusual expressive prosody can be a significant barrier to social acceptance, and it is inescapable that prosodic ability in autism is an area in need of more comprehensive and more systematic research.

References

AITKEN, K., PAPOUDI, D., ROBARTS, J. and TREVARTHEN, C., 1998, *Children with Autism, Diagnosis and Interventions to Meet their Needs* (London: Jessica Kingsley).

AMERICAN PSYCHIATRIC SOCIETY, 1980, *Diagnostic and Statistical Manual for Mental Disorders (DSM III)*, 3rd edn (Washington, DC: APS).

AMERICAN PSYCHIATRIC SOCIETY, 1987, *Diagnostic and Statistical Manual for Mental Disorders (DSM III-R)*, revd 3rd edn (Washington, DC: APS).

AMERICAN PSYCHIATRIC SOCIETY, 1994, *Diagnostic and Statistical Manual for Mental Disorders (DSM IV)* (Washington, DC: APS).

BALTAXE, C. A. M., 1984, Use of contrastive stress in normal, aphasic, and autistic children. *Journal of Speech and Hearing Research*, **27**, 97–105.

BALTAXE, C. A. M. and GUTHRIE, D., 1987, The use of primary sentence stress by normal, aphasic and autistic children. *Journal of Autism and Developmental Disorders*, **17**, 255–271.

BALTAXE, C. A. M. and SIMMONS, J. Q., 1985, Prosodic development in normal and autistic children. In E. Schopler and G. B. Mesibov (eds), *Communication Problems in Autism* (New York: Plenum), pp. 95–125.

BALTAXE, C. A. M., SIMMONS, J. Q. and ZEE, E., 1984, Intonation patterns in normal, autistic and aphasic children. In A. Cohen and M. van de Broecke (eds), *Proceedings of the 10th International Congress of Phonetic Sciences* (Dordrecht: Foris), pp. 713–718.

BARON-COHEN, S., 1995, *Mindblindness: An Essay on Autism and Theory of Mind* (Cambridge, MA: MIT Press).

BARON-COHEN, S. and STAUNTON, R., 1994, Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language*, **14**, 241–248.

BAUM, S. and PELL, M., 1999, The neural bases of prosody: insights from lesion studies and neuroimaging. *Aphasiology*, **13**, 581–608.

BOUCHER, J., LEWIS, V. and COLLINS, G. M., 1998, Familiar face and voice matching and recognition in children with autism. *Journal of Child Psychology and Psychiatry*, **39**, 171–181.

BOUCHER, J., LEWIS, V. and COLLINS, G. M., 2000, Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Psychology and Psychiatry and Allied Disciplines*, **41**, 847–857.

COUPER-KUHLEN, E., 1986, *An Introduction to English Prosody* (London: Arnold).

CRUTTENDEN, A., 1997, *Intonation*, 2nd edn (Cambridge: Cambridge University Press).

CRYSTAL, D., 1969, *Prosodic Systems and Intonation in English* (The Hague: Mouton).

ERWIN, R., VANLANCKER, D., GUTHRIE, D., SCHWAFEL, J., TANGUAY, P. and BUCHWALD, J. S., 1991, P3 responses to prosodic stimuli in adult autistic subjects. *Electroencephalography and Clinical Neurophysiology*, **80**, 561–571.

FAY, W. and SCHULER, A. l., 1980, *Emerging Language in Autistic Children* (London: Edward Arnold).

FINE, J., BARTOLUCCI, G., GINSBERG, G. and SZATMARI, P., 1991, The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, **32**, 771–782.

FLETCHER, P., 1985, *The Child's Learning of English* (Oxford: Blackwell).

FOSNOT, S. M. and JUN, S., 1999, Prosodic characteristics in children with stuttering or autism during reading and imitation. *Proceedings of the 14th International Congress of Phonetic Sciences*, pp. 1925–1928.

FRANKEL, F., SIMMONS III, J. Q. and RICHEY, E., 1987, Reward value of prosodic features of language for autistic, mentally retarded, and normal children. *Journal of Autism and Developmental Disorders*, **17**, 103–115.

GOLDMAN-EISLER, F., 1972, Pauses, clauses and sentences. *Language and Speech*, **15**, 103–113.

GREENFIELD, P. M., 1979, Informativeness, presupposition and semantic choice in single-word utterances. In E. Ochs and B. Schieffelin (eds), *Developmental Pragmatics* (New York: Academic Press), pp. 159–166.

HALLIDAY, M. A. K., 1967, *Intonation and Grammar in British English* (The Hague: Mouton).

HALLIDAY, M. A. K., 1970, *A Course in Spoken English: Intonation* (Oxford: Oxford University Press).

HARGROVE, P., 1997, Prosodic aspects of language impairment in children. *Topics in Language Disorders*, **17**, 76–83.

HARGROVE, P. and SHERAN, C., 1989, The use of stress by language-impaired children. *Journal of Communication Disorders*, **22**, 361–373.

HOBSON, P. R., 1986a, The autistic child's appraisal of expressions of emotion. *Journal of Child Psychology and Psychiatry*, **27**, 321–342.

HOBSON, P. R., 1986b, The autistic child's appraisal of expressions of emotion: a further study. *Journal of Child Psychology and Psychiatry*, **27**, 671–680.

HORNBY, P. and HASS, W., 1970, Use of contrastive stress by pre-school children. *Journal of Speech, Language and Hearing Research*, **13**, 395–399.

HOWELL, P. and KARIMA, K., 1991, Comparisons of prosodic properties between read and spontaneous speech material. *Speech Communication*, **10**, 163–169.

KENT, R. and READ, C., 2002, *Acoustic Analysis of Speech*, 2nd edn (San Diego, California: Singular, Thomas Learning).

KOWAL, S., WIESE, R. and O'CONNELL, D., 1983, The use of time in storytelling. *Language and Speech*, **26**, 377–392.

LOCAL, J. and WOOTTON, T., 1996, On being echolalic: an analysis of the interactional and phonetic aspects of an autistic's language. *York Papers in Linguistics*, **17**, 119–165.

MACWHINNEY, B. and BATES, E., 1978, Sentential devices for conveying givenness and newness: a cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behaviour*, **17**, 539–558.

McCALEB, P. and PRIZANT, B., 1985, Encoding of new versus old information by autistic children. *Journal of Speech and Hearing Disorders*, **50**, 226–230.

PACCIA, J. and CURCIO, F., 1982, Language processing and forms of immediate echolalia in autistic children. *Journal of Speech and Hearing Research*, **25**, 42–47.

PAUL, R., AUGUSTYN, A., KLIN, A., VOLKMAR, F. and COHEN, D., 2000, Grammatical and pragmatic prosody perception in high-functioning autism. Paper presented at the Symposium for Research in Child Language Disorders, Madison, WI, USA.

PEPPÉ, S., 1998, Investigating linguistic prosodic ability in adult speakers of English. PhD thesis, University College London.

PEPPÉ, S., MAXIM, J. and WELLS, B., 2000, Prosodic variation in southern British English. *Language and Speech*, **43**, 309–334.

PRICE, P., OSTENDORF, S., SHATTUCK-HUFNAGEL, S. and FONG, C., 1991, The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, **90**, 2956–2970.

PRIZANT, B. and DUCHAN, J., 1981, The functions of immediate echolalia in autistic children. *Journal of Speech and Hearing Disorders*, **46**, 241–249.

RAMBERG, C., EHLERS, S., NYDÉN, A., JOHANSSON, M. and GILLBERG, C., 1996, Language and pragmatic

functions in school-age children on the autism spectrum. *European Journal of Disorders of Communication*, **31**, 387–414.

RUTHERFORD, M. D., BARON-COHEN, S. and WHEELWRIGHT, S., 2002, Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders*, **32**, 189–194.

RUTTER, M., 1978, Diagnosis and definition of childhood autism. *Journal of Autism and Childhood Schizophrenia*, **8**, 139–161.

SCHOPLER, E., REICHLER, R. J. and RENNER, B. R., 1986, *The Childhood Autism Rating Scale (CARS)* (New York: Irvington).

SCHUETZE-COBURN, S., SHAPLEY, M. and WEBER, E., 1991, Units of intonation in discourse: a comparison of acoustic and auditory analyses. *Language and Speech*, **34**, 207–234.

SEDDOH, A. S., 2002, How discrete or independent are 'affective prosody' and 'linguistic prosody'? *Aphasiology*, **16**, 683–692.

SHEILDS, J., VARLEY, R., BROOKS, P. and SIMPSON, A., 1996, Hemispheric function in developmental language disorders and high-level autism. *Developmental Medicine and Child Neurology*, **38**, 473–486.

SHRIBERG, L. D., KWIATKOWSKI, J. and RASMUSSEN, C., 1990, *The Prosody-Voice Screening Profile* (Tucson: Communication Skill Builders).

SHRIBERG, L. D., PAUL, R., MCSWEENY, J. L., KLIN, A., COHEN, D. J. and VOLKMAR, F. R., 2001, Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger's Syndrome. *Journal of Speech, Language, and Hearing Research*, **44**, 1097–1115.

THURBER, C. and TAGER-FLUSBERG, H., 1993, Pauses in the narratives produced by autistic, mentally-retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental Disorders*, **23**, 309–322.

WANG, A., DAPRETTO, M., HARIRI, A., SIGMAN, M. and BROOKHEIMER, S., 2001, Processing affective and linguistic prosody in autism: an fMRI study. *Neuroimage*, **13**, S621–S621.

WELLS, B. and LOCAL, J., 1993, The sense of an ending: a case of prosodic delay. *Clinical Linguistics and Phonetics*, **4**, 59–73.

WIEMAN, L., 1976, Stress patterns of early child language. *Journal of Child Language*, **3**, 283–286.

WING, L. and GOULD, J., 1979, Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *Journal of Autism and Developmental Disorders*, **9**, 11–29.