Speech Science Research Centre

Receptive and Expressive Prosodic Ability in Children with High-Functioning Autism

Susan Peppé, Joanne McCann, Fiona E. Gibbon,
Anne O'Hare and Marion Rutherford

Working Paper WP-5

April 2006

Queen Margaret University College
EDINBURGH

QMUC Speech Science Research Centre Working Paper WP-5 (2006)
Series Editors: James M Scobbie, Ineke Mennen, Jocelynne Watson

THE QUEEN'S ANNIVERSARY PRIZES
2002

Update and Sourcing Information April 2006

This is a slightly altered version (v1.1) available May 2006, with minor typographical and pagination changes. This paper is available online in pdf format.

- 2006 onwards at http://www.qmuc.ac.uk/ssrc

Author Contact details:

- 2006 onwards at speppe@qmuc.ac.uk

Subsequent publication & presentation details:

- This paper has been reviewed twice and subsequently revised for the *Journal of Speech Language and Hearing Research.* Reference should be made to the JSLHR version when it appears.

QMUC Speech Science Research Centre Working Paper WP-5 (2006)
Series Editors: James M Scobbie, Ineke Mennen, Jocelynne Watson

Receptive and Expressive Prosodic Ability in Children with High-Functioning Autism

Sue Peppé*, Joanne McCann,* Fiona Gibbon,*

Anne O'Hare°† and Marion Rutherford*°

**Queen Margaret University College, Edinburgh*

°Royal Hospital for Sick Children, Edinburgh

† University of Edinburgh

Abstract

This study aimed to identify the nature and extent of receptive and expressive prosodic deficits in children with high-functioning autism. In a data-based group study, 31 children with high-functioning autism (HFA, excluding Asperger's syndrome) and 72 typically developing controls matched on verbal mental age completed a prosody assessment procedure (PEPS-C). Children with HFA performed significantly less well than controls on eleven out of twelve prosody tasks ($p < .005$). Receptive prosodic skills showed strong correlation ($p < .01$) with verbal mental age in both groups, as did, to a lesser extent, expressive prosodic skills. Receptive prosodic scores also correlated with expressive prosody scores, particularly in grammatical prosodic functions (turnend and prosodic phrasing/ chunking). Prosodic development in the HFA group appeared to be delayed in many aspects of prosody and deviant in some (e.g. accent tended to be placed early in focus tasks and Same items were often perceived as Different in auditory discrimination tasks). The study demonstrates that receptive prosodic deficit, expressive prosodic skills, and language development are closely associated in the condition of autism. Receptive prosodic skills would be an appropriate focus for clinical intervention, and further investigation of atypical expressive prosody and the relationship between prosody and social skills is warranted.

Key Words: high-functioning autism, prosody, intonation, language, assessment.

1. Introduction

Prosody in Autism

An association between autism and prosody has long been suspected because children with autism frequently have expressive prosody that is atypical and not like that found in other disorders (Baltaxe & Simmons, 1985; Fay & Schuler, 1980). If there is a link, then either some aspect of autism may give rise to at least some aspects of prosodic disorder, or vice versa. By establishing whether there is a link, and, if so, the nature of the link, it may be possible to increase understanding of the aetiology and nature of autism and ameliorate some of the aspects of impaired communication associated with it. Previous studies in this area (reviewed in McCann & Peppé, 2003) have focused on expressive prosody, while little attention has been paid to receptive skills; the focus of this paper is on receptive as well as expressive prosodic skills and the relationship between them.

Autism

Autism is a pervasive developmental disorder that begins in the first 36 months of life (American Psychiatric Society DSM-IV, 1994; ICD-10, World Health Organisation, 1993) and has been defined as a triad of impairment: atypical development in reciprocal social interaction; atypical communication; and restricted, stereotyped and repetitive behaviours (Wing & Gould, 1979). It is a spectrum disorder: individuals may be nonverbal with learning difficulties (having low-functioning or classic autism), while at the upper end of the spectrum individuals are cognitively within normal limits and may be classed as having high-functioning autism (HFA) or Asperger's Syndrome (AS), the distinction between HFA and AS being controversial (Macintosh & Dissayanake, 2004). The most recent diagnostic criteria hold that individuals with AS, unlike those with HFA, do not demonstrate a general preschool language delay, and indeed may be highly articulate. From this ranking it can be seen that the degree of severity of autism is strongly, not to say circularly, connected with degree of language ability (at least as regards expression/output). In this study we limit our investigation to individuals with HFA but not AS, i.e. to individuals in whom language delay (as well as other features of autism) was apparent before they entered school.

There are reasons apart from the frequent presence of atypical expressive prosody to think that problematic prosodic ability may be particularly associated with autism. People with autistic spectrum disorders have difficulties with social relationships; difficulties which may have their roots in impaired comprehension of communicative messages. They also tend to be literal in their interpretation of language (DSM-IV), perhaps because they have difficulty in generalising from specific events and in drawing inferences from experience (weak central coherence, Happé, 1997): this may result in a failure to access some of the non-literal aspects of language that enable social interaction. Prosody is such an aspect: although people can be wholly influenced by its effects, they tend not to recognize overtly that it is prosody that mediates their reactions. In support of this, it has been observed that speakers often do not correct their own prosodic errors (Cutler, 1980); this suggests that frequently prosody does not impinge on speakers as having a role in communication. Furthermore, it appears that misleading prosody generally goes uncorrected as language develops, whereas a care giver often corrects a child's erroneous pronunciation, grammar or lexical choice.

A skill such as prosody which unobtrusively conveys emotional and pragmatic aspects of speech may be particularly vulnerable in children with autism, perhaps even one of the causes of their characteristic social deficits. Children with autism who do not have learning difficulties may, however, be capable of increasing their prosodic awareness and ability. This in turn may reduce some of the communication barriers and increase their social reciprocity

In the research literature, numerous adjectives are used to describe atypical expressive prosody in autism , e.g. "dull", "wooden", "singsong", "robotic", "stilted", "over precise" and "bizarre" (Baltaxe & Simmons, 1985; Fay & Schuler, 1980): terms which perhaps reflect perceived characteristics of autism more than acoustic features. The fact that adjectives as opposed as “monotonous” and “exaggerated”, “slow” and “fast” (Baron-Cohen & Staunton, 1994) can be used to describe this atypicality suggests a wide variation in either the perception of atypical expressive prosody or in the prosody itself. A characterisation of the acoustic features of atypical expressive prosody is lacking in the literature, and is beyond the scope of this paper. Instead, we focus on the ability of children with autism to make use of prosody to infer and express meanings.

If the terms used to describe the atypical expressive prosody of autism are many and various, the terminology and range of prosody is also diverse, and prone to confusion.

Prosodic functions and forms

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 (Cruttenden, 1997), but what is included in the domain of “prosody” can vary greatly. Because of this, the scope of an investigation into prosody in autism tends to vary from one study to another.

The term ‘prosody’ is used to describe suprasegmental characteristics of speech. These derive from variations in the duration, amplitude and fundamental frequency of speech-sounds, and comprise the acoustic features of communicative functions conveyed by prosody, but not the functions themselves. These features, including both what may be described as “intonation” (pitch-variations) and suprasegmental breaks in speech (pauses), are included in the umbrella term “prosody” in the current study. “Prosody” (with initial capital) is used in the prosody assessment procedure to denote the suprasegmental characteristics of prosody required for the accentual and grouping functions associated with multi-word phrases, and is distinct from “Intonation”, which concerns only the variations (ones of pitch-height, span and contour) required to convey different meanings on single words.

In the literature, there is often confusion in prosodic terminology, leading to circularity or failure to address research aims. For example, in a recent study (Paul et al, 2005), one of the prosodic forms was named “phrasing” and defined as “the pattern of rate and pausing within utterances” (p.207). One of the authors’ valid and interesting aims was to see if “phrasing” could be used in pragmatic/affective functions as well as grammatical ones. This would have indicated whether there was disorder in both functions: if so, this might indicate a problem at the form level. However, the task devised to exemplify the grammatical function played on “the ability to use *pauses* to signal referential vs. non-referential phrases within sentences” (p.214, my italics), whereas the one for the pragmatic/affective function (the distinction between calm vs. excited speech), produced a role for *rate*: subjects were observed to cue in to the rate difference and “used it exclusively to solve the problems posed in this task” (p.214). The connection of a common form between the two tasks was therefore lost, and it seems likely that this was at least partly due to confusion as to what the terms “phrasing” and “rate” referred to. “Phrasing” was to have been achieved by both rate and pausing, but the tasks separated them. It appears that “rate” was taken to refer to both final syllable lengthening, i.e. relative syllable rate, a crucial indicator in prosodic phrasing (Scott, 1982; Katz, Beach, Jenouri & Verma, 1996), and tempo, or overall speech-rate, important for affective distinctions. Such a blurring of distinctions is by no means an isolated occurrence in the literature involving prosody research, and highlights the need for better definition of the forms and functions of prosody.

Communicative functions in which prosody plays an important part have been categorised (Roach, 2000) as grammatical, pragmatic, and affective, although as far as prosody is concerned these categories are not mutually exclusive. The following paragraphs show how the prosody test used in this study (the PEPS-C, see Appendices I and II) takes these functional distinctions into account.

Grammatical functions of prosody include the segmenting of utterances into prosodic phrases; the ends of phrases are signalled by a number of prosodic factors (forms) including: pause (Butcher, 1981); lengthening of the final syllable (Scott, 1982); and the presence of nuclear tone (Cruttenden, 1997), located at or near the end of the utterance. Such boundary markers have been collectively designated ‘boundary tone’ (Pierrehumbert, 1987). For example, in utterances such as "chocolate cake and cream", or "Ellen the dentist is here" there will be boundary markers on the final word. There is, however, potential ambiguity according to whether or not there are also grammatical boundaries within the utterance ("chocolate, cake and cream" or "chocolate-cake and cream"; "Ellen, the dentist is here" or "Ellen, the dentist, is here") and the presence or absence of prosodic boundary markers within the utterances can disambiguate them. In the PEPS-C, this segmentation function is designated Chunking.

In the absence of any need for particular emphasis, the focus of an utterance is said to be "broad" as in the neutral utterance "I wanted some socks". When particular importance or contrast is (pragmatically) required on one word or syllable, the accent is placed on that word/syllable and the focus of the utterance is described as "narrow". An example is the utterance "I wanted BLUE socks" where "blue" is accented to suggest that this is an important word, probably contrasting with a previously mentioned colour. Narrow focus can be located at any point within an utterance. Accent is manifested by the relative prominence (boosted pitch, extra length and/or loudness) of the relevant syllable (Fry, 1958). The function designated Focus in the PEPS-C refers to narrow or contrastive focus.

Indication of sentence-type by the use of intonation, sometimes called inflection, can be considered as an interactional or grammatical function: for example, a sharply rising intonation at the end of an utterance implies that some response is required, while a falling intonation usually suggests finality (Cruttenden, 1997). This function is referred to as Turnend in the PEPS-C.

Finally, affective functions include the 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 (Mozziconacci, 1998), which tend to apply to an entire utterance rather than a few syllables of it. 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 (Banse & Scherer, 1996). Tasks addressing this function in the PEPS-C are designated Affect.

Research into Prosody in Autism

A recent paper reviewed the sixteen studies that were available and relevant to the topic of prosody in autism (McCann & Peppé, 2003). The main prosodic topics of research were the placement of accent and boundary and production of intonation patterns distinguishing utterance type. Accent (both its placement and its realisation) was found to be disordered in individuals with autism in several studies (Shriberg, Paul, McSweeny, Klin, Cohen, & Volkmar, 2001; Fine, Bartolucci, Ginsberg, & Szatmari, 1991; Baltaxe & Guthrie, 1987; Baltaxe & Simmons, 1985), and intonation patterns distinguishing utterance type were also problematic. The picture that emerged was inconclusive, however. Findings were sometimes contradictory; for example, one study of intonation production found that this was disordered for individuals with autism (Fosnot & Jun, 1999) and another that individuals with autism did not differ from typically developing controls (Baltaxe & Simmons, 1984). McCann & Peppé (2003) suggested two main reasons for this. First, diagnosis of autism was not consistent for all study participants. There was also a discrepancy of investigative methods: various procedures were used to assess prosody, and studies on the whole involved few participants with autism and were lacking in control data. Furthermore, some relevant aspects of prosody (e.g. receptive abilities) were seriously under researched. Given the unusual quality of expressive prosody in autism, it is not surprising that researchers' attention has been directed towards prosodic expression. It would however be logical to suspect that the cause of atypical expressive prosody might lie at least partly in reduced ability to process prosody or to perceive it as meaningful, and this notion is the reason for investigating receptive and expressive skills in comparable tasks.

A more recent paper (Paul et al., 2005) investigates perception and production of prosody in children with autism. The study is considered in some detail because its procedures closely resemble those of the current one: Paul et al used six tasks eliciting expressive data and six testing receptive skills in three different functional aspects of prosody and three aspects of prosodic form. Paul et al.'s participants were different, however, comprising 27 individuals with high-functioning autism /Asperger's syndrome /Pervasive Development Disorders Not Otherwise Specified and 13 typically developing controls, both groups older than those in the current study (mean chronological age of 16+ years). Results suggest little difference between the performances of Paul et al's two groups on most of the tasks, but the authors conclude that this is less because of a lack of difference in prosodic ability and more because of weaknesses of task design. Ceiling effects were apparent in both groups on some tasks; where this occurred in perception tasks, the authors suggest that prosodic differences may have been too apparent, and the tasks therefore too easy. On the other hand, typically developing 16-year-olds did not achieve near-ceiling scores in some of the perception tasks (including the two concerning stress); this suggests a lack of ecological validity in the tasks, because such skills are likely to have been acquired at an earlier age. The authors also point out that there were disadvantages in having teenagers as participants, as they found some of the tasks (e.g. the need to produce 'motherese') embarrassing. An offsetting advantage of having older participants was that reading ability was well established in participants of both groups and required for all the production tasks; this however begs the question of how expressive prosody would be realised if, as in conversation, the lexis had to be supplied at the same time as the prosody. Another reason for lack of score differences could have been that the age and clinical diagnosis of the participants was wide ranging. There is also the possibility that a control group numbering 13 is perhaps too small to allow for normal variability. Despite these caveats, the study found, as in the studies previously cited, that production of stress, whether pragmatically (for purposes of contrast) or for grammatico-lexical distinction, was disordered in the group with autistic spectrum disorders.

Aims of Current Study

From the foregoing review of the literature it appears that the general perception that prosody is disordered in autism has not been fully substantiated by research; but on balance faulty methodology is taken to be more responsible for lack of substantiation than a faulty premise. The present study aimed to avoid as many as possible of the pitfalls of previous studies, but the hypothesis remains the same: that prosodic ability is lower in children with high-functioning autism than in typically-developing children. In addition, it was thought that receptive prosody in particular would merit examination, and that it would be beneficial to analyse any between-group differences in terms of deviance or delay.

2. Method

One characteristic of previous studies was that they tended to investigate single aspects of prosody in heterogeneous groups of subjects. Our study instead opted for comprehensive examination of prosody in a narrowly defined group of children with autism.

Participants

Thirty-one children with HFA took part: 24 boys, 7 girls, age range 6;1 to 13;6 (*M* = 9;10, *SD* = 2.3); and 72 typically developing children: 54 boys, 18 girls, age range 4;10 to 11;8 (*M* = 6;10, *SD* =1.5). Criteria for inclusion in the group with HFA were as follows: a diagnosis of autism (but not Asperger's Syndrome) in preschool years; ages of 6 to 13 years; verbal mental age > 4;0 years; reports of nonverbal ability within the normal range; no significant binaural hearing loss, no significant visual impairment, no major physical disability and no structural abnormality of the vocal tract; English the first language and the language of the home; and residence in the Edinburgh area of Scotland for at least 3 years.

Inclusion criteria for the children with typical development (TD) were that they should have a verbal mental age > 4;0 years and socioeconomic status similar to that of the children with high-functioning autism. To achieve the latter criterion, the Carstairs Index (McLoone, 2000), a classification of deprivation, was used to identify Edinburgh state primary schools in which the typically developing children were from the same postcode areas as the children with autism. The group with TD was drawn from these schools and had no significant hearing loss or visual impairment, no major physical disability or structural abnormality of the vocal tract; English was the first language and the language of the home; and they had been resident in the Edinburgh area of Scotland for at least 3 years. 76 controls were recruited, to ensure a likelihood that at least one child would be matched with each of the children with HFA both verbally and socioeconomically: verbal mental age could not be ascertained in advance of recruitment as no clinically recognised measure is conducted as part of a typically developing child's education. Of these, 72 met verbal mental age criteria and were included in the study.

Autism Diagnosis

The diagnosis of autism, made by multidisciplinary assessment, was based on ICD 10 (World Health Organisation, 1993) or DSM-IV (pervasive developmental disorders: American Psychiatric Society, 1994), and a range of other autism assessment tools: the Childhood Autism Rating Scale (DiLalla & Rogers, 1994), Gilliam Autism Rating Scale (Gilliam, 1995), and Autism Diagnostic Observation Schedule (Lord, Risi, Lambrecht, Cook, Leventhal, & DiLavore, 2000). Diagnosis also took account of clinical observations with regard to communication, reciprocal social interaction and repetitive behaviours, noting the children's ability to comprehend, imitate, use and attend to language; to interact socially; and to play appropriately with toys. Children previously diagnosed with Asperger's syndrome were not selected.

Procedures

Findings reported in this study are from two procedures: Profiling Elements of Prosodic Systems in Children (PEPS-C: Peppé & McCann, 2003), which measures receptive and expressive prosody, and the British Picture Vocabulary Scales II (BPVS II: Dunn, Dunn, Whetton, & Burley, 1997), which measures receptive vocabulary and is a standard measure of verbal mental age used in other studies of children with autism (e.g. Baron-Cohen, Leslie, & Frith, 1985, Thurber & Tager-Flusberg, 1993). The group with HFA also completed a battery of language assessments as part of a larger research project; results of these are reported in a separate paper (McCann, Peppé, Gibbon, O'Hare, & Rutherford, under review).

Prosody Assessment

PEPS-C is a relatively new prosody assessment procedure. A manual version has been used in previous studies (Wells & Peppé, 2003; Wells, Peppé, & Goulandris, 2004). The computerized Scottish version is described in Peppé and McCann (2003). Details of tasks, instructions, scoring procedures and task items are given in the Appendices. The procedure aims to evaluate prosodic skills at the level of form processing (auditory discrimination and ability to produce prosodic variation) and of communicative function (pragmatic/interactional, affective, and grammatical/linguistic) in two communicative modes, receptive (input) and expressive (output).

Function Tasks

In parallel input and output tasks, PEPS-C assesses most of the aspects covered in the literature on prosody in autism (see McCann & Peppé, 2003) by sampling prosodic functions grouped under four headings: Turnend (indicating whether an utterance requires an answer or not), Affect (indicating liking or reservation, in this test with respect to certain foods), Chunking (signalling prosodic phrase boundaries), and Focus (emphasising one word in an utterance for contrastive accent).

Reading ability was not required, all responses being elicited by pictures. Items were selected to be culturally inoffensive, unlikely to elicit responses predictable from non-prosodic factors, and likely to be familiar to young children; a vocabulary check was run, with pictures being shown to and named by each child. Segmentally, the items were easy to pronounce, and monosyllables with non-sonorant segments were excluded. For the input tasks, items were agreed by at least two judges to indicate the meaning unambiguously but not exaggeratedly.

To avoid a heavy demand on auditory memory, only two response choices were offered for each item in input tasks. It was therefore necessary to include at least 16 items per task in order to have a reasonable number of non-chance scores (> 11 and < 5); the number of items was no higher than 16 in order to keep the demands on attention as low as possible. Because of the relatively wide chance-band, children were deemed to have reached competence level in a task if their score was at least 12 (75%), rather than half marks. In output tasks there was also an element of producing the right answer by chance, so the competence level was similarly set at 75%.

Form Tasks

In PEPS-C, variations in pitch range and tone direction with durational factors, which are used for conveying Affect and Turnend, are subsumed, as previously mentioned, under the term "Intonation", while the exponents of Chunking and Focus - variation in syllable duration, syllable prominence, and pause factors – are described as "Prosody". Input tasks, testing auditory discrimination, follow a Same/Different format; stimuli are from laryngograph signals (i.e. suprasegmental information only, without lexis) recorded simultaneously with the function task stimuli. For the Intonation input task, equal numbers of stimuli from the Affect and Turnend function tasks were used; for the Prosody input task, equal numbers of Chunking and Focus stimuli. Output tasks sought to elicit an Intonation and Prosody repertoire by requiring participants to imitate stimuli that were similar to, and extended the range of, stimuli used in the input function tasks.

Assessment Conditions

Children with HFA were interviewed by a certified speech and language therapist (the second author) in a clinic room designed for recording speech and language therapy sessions; parents and siblings could be present if wished. Typically developing children with typical development were mainly interviewed in a quiet room in their schools. Both groups were interviewed by the same person using the same equipment. Children were seated approximately 18 inches from the computer screen, and were free to leave the assessment whenever they wished. Children with HFA were accompanied by a caregiver able to say if necessary whether the child wished to cease participation.

The prosody tasks lasted between 40 minutes and an hour, according to variability of children's age, ability and attention spans. For the HFA group, two one-hour sessions were necessary for the complete battery of assessments. The interviewer explained each prosody task and administered two practice items per task. If the children failed the practice tasks they were readministered; if these failed, the task was aborted. Sessions were recorded on DAT tape and CD with computer backup.

The PEPS-C program delivers pictures successively on a laptop screen both as stimuli for expressive utterances and as response choices to auditory stimuli played by the computer. A list of the pictures is given in Appendix II. In expressive tasks, the tester sat out of line of sight of the screen, to ensure neutrality of judgment. The children responded to the picture stimulus in their own time, with the stimulus constantly visible. After the child's response, the next stimulus appeared, in randomised order, in response to a key press by the tester as she made her judgment on the child's response. Judgments were made without the use of headphones, to allow the tester to evaluate them as they might be perceived in a natural hearing environment. In receptive tasks, the children made judgments by clicking on the half of the screen that showed the relevant response choice; this click prompted the next stimulus, in a fixed order. For these tasks, the tester sat so as to ensure that the child's mouse click was not ambiguously close to the border between the two halves of the screen. In certain circumstances, the tester replayed the item: where children had been distracted, or had not heard the item (because of ambient noise, for example); or if children said they had made a wrong mouse click. Instructions and scoring procedures are given in Appendix I, and a demonstration of the program can be found on the website: http://www.qmuc.ac.uk/ssrc/prosodyinasd/

Statistical Methods

Analysis of covariance was used to establish significance of group differences on task scores, and to establish the role of chronological age. Although the HFA group showed greater variance in some of their scores, parametric methods were deemed to be sufficiently robust to justify use of the *F*-test in this analysis. The relationships between the various measures were established using Pearson's correlations. Post-hoc analyses (*t*-tests) were used to investigate significance of differences in error types. Because of the large number of tests, difference was considered significant if $p < .01$.

3. Results

Differences of Age, Sex and Deprivation Category

As expected, the groups did not differ significantly ($t(101) = -1.22$, $p = .302$) on the verbal mental age measure; the mean BPVS II age equivalent for the HFA group was 7.09 ($n = 31$, $SD = 2.01$) and for the TD group 7.53 ($n = 72$, $SD = 1.55$). Again as expected, verbal ability was significantly different ($t(40.18) = 8.598$, $p < .001$), with the mean BPVS-II standard score 81.6 ($SD = 15.6$) in the HFA group and 107.5 ($SD = 9.6$) in the TD group.

Chronological age was significantly different ($t(101) = 7.919$, $p < .001$), the HFA group having a mean age of 9;10, the TD group 6;10. No significant group difference was found on either social deprivation category or sex.

Prosody Task Scores

The mean percentage prosody scores of the HFA group were compared with those of the TD group; results are shown in Table 1. The TD group have mean scores above competence level (but not at ceiling) in 7 out of the 12 tasks: Affect Input, Affect Output, Focus Output, and the four form tasks; they were approaching competence on two output tasks (Chunking and Turnend) and below 70% on the remaining three input tasks (Chunking, Focus, and Turnend). The HFA group's mean scores were approaching competence on one task (Affect Input), and below 70% on the other eleven. Participant numbers ranged between 29 and 31 for the HFA group and between 71 and 72 for the TD group.

Table 1. *Analysis of covariance (F) between groups on PEPS-C tasks (in = input/receptive, out =output/expressive) showing ranges, percentage mean scores (M), standard deviations (SD) and probability (p).*

Task:	Turnend in		Turnend out		Affect in		Affect out		Intonation in		Intonation out	
Group:	HFA	TD	HFA	TD	HFA	TD	HFA	TD	HFA	TD	HFA	TD
n	31	72	30	72	31	72	30	72	31	72	30	72
Range:min	18.7	31.3	43.7	31.3	31.3	50.0	18.7	31.3	25.0	31.3	9.4	34.4
max	100	100	100	100	100	100	100	100	100	100	93.8	100
M	65.9	64.8	68.1	74.2	71.2	84.5	63.3	79.4	68.8	80.1	64.7	79.9
SD	21.4	18.1	21.8	18.3	21.6	11.4	26.3	19.2	22	17.4	20.6	18.0
$p(df\ 1{,}97)$	p =.001		p <.001		p <.001		p =.003		p <.001		p =.001	
F	11.73		18.17		25.95		9.17		28.93		11.05	

Task:	Chunking in		Chunking out		Focus in		Focus out		Prosody in		Prosody out	
Group:	HFA	TD	HFA	TD	HFA	TD	HFA	TD	HFA	TD	HFA	TD
n	31	72	30	72	30	71	29	71	31	72	30	72
Range:min	43.7	37.5	18.7	37.5	31.3	37.5	18.7	56.3	31.3	43.8	34.4	37.5
max	93.8	93.8	100	100	100	100	100	100	100	100	100	100
M	67.5	69.0	66.5	71.2	59.6	65.9	61.6	84.0	63.5	79.0	65.7	85.4
SD	15.7	15.6	26.4	11.8	19.0	19.1	26.4	15.0	23.1	13.3	22.0	11.4
$p(df\ 1{,}97)$	p =.008		p =.623		p <.001		p <.001		p <.001		p <.001	
F	7.42		.243		18.34		44.07		48.76		30.59	

Analyses of covariance with chronological age as the covariate factor (also shown in Table 1) found significant group differences (*df* 1, 97, p<.001) on seven tasks (Affect Input, Turnend Output, the Focus tasks, Intonation Input, and the Prosody tasks); and at the .01 level on four (Affect Output, Turnend Input, Chunking Input and Intonation Output), with no significant difference on Chunking Output.

Reliability

Only the output scores were subject to rater variation. For a measure of intrarater reliability, the tester rejudged 10% of the data after a lapse of at least six months by listening to the CD recording of the children's responses, using headphones (Sennheiser eh2200) and comparing original with later judgments. Rejudged data consisted of the complete set of output scores of four children with HFA and of seven with TD. Cronbach's alpha was used to calculate the intra-class correlation co-efficient, which was at a mean of .91 (p < .001), minimum .80 (Chunking), maximum .95 (Prosody). The earlier set of task scores was used in calculations of the children's performance and in interrater comparisons.

As an interrater measure, we used the tester's judgments and those of an experienced speech and language therapist (the fifth author) who was inexperienced in using the PEPS-C; both had Edinburgh accents similar to those of the testees. The fifth author listened to 10% of the tasks under similar conditions. Mean Cronbach's alpha was at .82 (p <.001), minimum .78 (Chunking Task), maximum .98 (Intonation).

<u>Practice Effects</u>

Paired samples t-tests were used to compare performance on the first five items of each task with the last five. In the HFA group, no differences were significant at the .01 level, suggesting children neither learned nor tired during the tasks. In the TD group, there was evidence of tiring in two tasks (Intonation Output, $t(72) = 3.6$, $p = .001$; Prosody Input, $t(72) = 2.99$, $p = .004$).

<u>Acquisition of Prosody and Language Skills</u>

Table 2 shows the degree to which each of the skills measured by PEPS-C correlates with verbal mental age in typically developing Edinburgh children and in the HFA group. Mean scores of 28 adults are also shown, to indicate that children may reasonably be expected to acquire these skills eventually. The adults were tested in conditions similar to those used for testing children; two interviewers shared the testing. The adults either had Scottish accents or had been resident in Scotland for at least three years.

Table 2. *Correlation (Pearson's) of PEPS-C scores with children's verbal mental age (VMA), showing overall and individual receptive (in) and expressive (out) task mean scores for the TD group, and the HFA group; and mean scores of 28 neurotypical adults.*

	Overall		Turnend		Affect		Chunking		Focus		Intonation		Prosody	
	in	out	in	out	in	out	in	out	in	out	in	out	in	out
TD *n*	72	72	72	72	72	72	72	72	71	71	72	72	72	72
M	73.9	79.0	64.8	74.2	84.5	79.4	69.0	71.2	65.9	84.0	80.1	79.9	79.0	85.4
VMA: *r*	.691	.502	.694	.528	.435	.305	.468	-.005	.419	.073	.443	.422	.446	.524
p	.<001	.<001	.<001	.<001	.<001	.<001	.<001	ns	.<001	ns	.<001	.<001	.<001	.<001
HFA *n*	31	30	31	30	31	30	31	30	30	29	31	30	31	30
M	66.1	65.0	65.9	68.1	71.2	63.3	67.5	66.5	59.6	61.6	68.8	64.7	63.5	65.7
VMA: *r*	.773	.310	.628	.550	.583	.069	.544	.091	.572	.412	.585	.039	.467	.085
p	.<001	<.096	<.001	=.002	=.001	ns	=.002	ns	=.001	=.027	=.001	ns	=.008	ns
Adult *n*	28	28	28	28	28	28	28	28	27	26	28	28	28	27
M	93.1	91.3	92.1	88.4	93.6	94.1	91.3	93.6	92.1	91.7	94.1	92.1	95.6	88.0

In the TD group, all but two of the prosody tasks (Chunking Output and Focus Output) show highly significant correlations with verbal mental age. Receptive prosody scores, in particular, correlate highly ($p < .001$) and there is a trend towards better expressive scores. A complete investigation of whether prosodic skills improve with age in children with autism would require a longitudinal study of the development of prosodic skills within individual children with autism.

<u>Deviance in Development of Prosodic Ability</u>

Deviance in prosodic development was defined as a significant difference between the groups on errors. Numbers and direction of response options were compared, in order to find patterns of errors. Some differences in error patterns were discernible:

<u>Form Tasks</u>

In Intonation input tasks, the HFA group were significantly more likely than the TD group to judge Same items as different ($t = 3.74$, $p < .001$) (see Appendix II). In Prosody input tasks, this difference was even greater ($t = 6.52$, $p < .001$).

<u>Turnend</u>

In the input task, there was no significant difference in judgment preferences between groups, but many of the HFA group (12.9% as opposed to 2.7% of the TD group) judged all question-type stimuli as statements and none as questions (see Appendix II). In the output task, statement responses from the HFA group were significantly more likely than those of the TD group to be judged questioning ($t = 2.78$, $p = .006$) or ambiguous ($t = 2.34$, $p = .019$).

<u>Affect</u>

In the input task, the HFA group made more errors of both types than the TD group: in failing to discern liking a food item (see Appendix II) ($t = 3.97$, $p < .001$), and in failing to discern disliking ($t = 2.93$, $p = .003$). In the output task, their liking responses tended to be judged as disliking ($t = 3.27$, $p = .001$) and disliking as liking ($t = 3.61$, $p < .001$). Judgments of ambiguity were more frequent than for those in the TD group, both for liking ($t = 3.57$, $p< .001$) and for disliking ($t = 2.33$, $p= .02$).

<u>Chunking</u>

In the input task there were no discernible error patterns. In the output task, children with HFA were slightly more likely ($t = 2.4$, $p = .016$) than the TD group to fail to make the prosodic breaks necessary to disambiguate utterances.

<u>Focus</u>

In the input task, the HFA group made significantly more errors than the TD group when the accent was early, i.e. on the colour (see Appendix II) than when it was later (on the animal) ($t = 4.36$, $p < .001$). In the output task, their responses showed a significantly greater tendency than the TD group to place accent on the colour ($t = 3.5$, $p < .001$). HFA responses of both types were more ambiguous than those of the TD group (accent on colour: $t = 3.43$, $p = .001$; accent on animal: $t = 5.2$, $p< .001$)

<u>Receptive and Expressive Scores</u>

Correlations between receptive and expressive scores in the HFA and TD groups are shown in Table 3.

Table 3.
Correlations between Receptive and Expressive Scores for Overall and Individual PEPS-C Tasks for HFA and TD Groups, showing r-values and probability.

	HFA						
	Mean	Turnend	Affect	Chunking	Focus	Intonation	Prosody
r	.557	.740	.382	.615	.317	-.002	.128
p	.001	<.001	.019	<.001	.047	ns	ns

	TD						
	Mean	Turnend	Affect	Chunking	Focus	Intonation	Prosody
r	.732	.605	.583	.095	.260	.459	.557
p	<.001	<.001	<.001	ns	.029	<.001	<.001

Strong correlations between overall mean receptive and expressive prosody scores were apparent in the TD group ($p < .001$) and in the HFA group ($p = .001$). For the TD group, there was weak or nonsignificant correlation between input and output tasks for Chunking ($r = .095$) and Focus ($r = .260$) but strong correlation for the tasks depending on intonational differences, Turnend ($r = .605$, $p < .001$) and Affect ($r = .583$, $p < .001$). For the HFA group, there seemed to be stronger correlation between receptive and expressive scores in the tasks that might be described as associated with grammar: Turn-end ($r = .740$, $p < .001$) and Chunking ($r = .615$, $p < .001$) than in tasks more associated with pragmatic/affective functions (Affect ($r = .382$, $p = .019$) and Focus ($r = .317$, $p = .047$). Within form tasks, findings for the HFA group showed that discrimination (input ability) did not correlate with the ability to imitate utterances (output ability); for the TD group, by contrast, correlation in form tasks was highly significant (Intonation $r = .459$, $p < .001$; Prosody $r = .557$, $p < .001$).

4. Discussion

The correlation of chronological age with task performance in the typically developing group (Table 1) supports evidence from the earlier studies already cited that the prosodic skills tested in PEPS-C continue to develop during childhood. Two tasks (Chunking Output and Focus Output) showed no development; but apparently for different reasons. On the Focus Output task, the youngest children in the typically developing group produced scores well above competence level, suggesting they already possessed this skill and that it develops little after the age of 5. On the Chunking Output task, scores are just below competence level throughout the age range, but in the Wells et al. study, carried out in London using an earlier version of the PEPS-C test, the youngest children were at competence level on this task. Edinburgh children's scores may have been reduced by the introduction of items involving colours ('colour items', see Appendix II): in Wells et al.'s study only food items were used. It was judged that the colour items would provide interesting extra data: being ad hoc combinations of colours, they are more truly a reflection of grammatical phrasing effected by prosody than the food items, in which prosody conveys lexical distinctions. A *t*-test showed that the typically developing group scored less well on colour items than on food items (t 15.879, $p < .001$). A study of prosodic production in children aged 5 and 7 and adults (Katz et al, 1996) involved a similar task, grouping pink, white and green blocks (pink&green and white vs. pink and green&white); the children did not make chunking distinctions, whereas the adults did. In Table 2 it is shown that the adults' mean score was 95% on the task that included both colour and food items, so it must be concluded that the skill needed to produce colour items intelligibly is acquired after the age range tested in this study. As expected, prosody scores correlated highly with verbal mental age on all except the same two tasks, Focus Output and Chunking Output.

For the HFA group, receptive prosody scores correlated only weakly with chronological age ($r = .378$, $p = .036$), but strongly with verbal mental age ($r = .773$, $p < .001$), while expressive scores were not significantly correlated with either. The correlation between PEPS-C receptive task scores and verbal mental age suggests a stronger association between verbal ability and receptive prosody skills in HFA than between receptive prosody and chronological age, but in general a picture of delay. Expressive skills do not however correlate strongly with age, chronological or verbal, in the HFA group, and the development in this respect of the children in this group would appear to be deviant.

Auditory Processing in Form Tasks

Table 1 showed that children with autism performed significantly less well than typically developing peers on input form tasks (auditory discrimination). The tendency to perceive difference in Same items was seen as deviant, because simply not perceiving differences would produce the opposite tendency (perceiving Different items as Same). There is no consistency of item in the Same items perceived as differing by the HFA group, and therefore across-item variation is unlikely to account for the group difference. One possibility is that there is a general auditory memory problem in children with HFA, but this would not produce a significant tendency to judge Same items as Different. Another possibility is that if ambient noise was present during one half of a Same stimulus, the children with HFA may have perceived the noise as part of the stimulus, whereas the typically developing children were able to select only the stimulus for comparison. Other reasons may be postulated, but it is possible that prosodic discrimination ability may be deviant in autism.

Prosodic Functions

The Turnend tasks show that children with autism had a tendency to judge offering (questioning) stimuli as reading (stating) rather than to judge statements as questions, and to sound questioning when a statement was required, as opposed to the other way round.. Incomplete appreciation of the concept of questioning might predispose a speaker to produce statements in all situations, but this would not explain the tendency to sound questioning when a statement is required.

In the Affect tasks errors were fairly evenly distributed across both options, suggesting a more equal difficulty. Observation showed that children were inclined to judge input stimuli according to their own preferences, despite reminders that they were to consider the preferences of the speaker on the computer. The high degree of ambiguity for both options in the output task suggests that responses bore little relation to the speaker's preferences.

Children with HFA differ least from the typically developing group in the Chunking tasks: they were marginally more likely to fail to make prosodic breaks where necessary, but there were no appreciable differences in preferences for response choices on the input task, nor in the number of responses judged as ambiguous in the output task. This is perhaps some evidence in favour of the hypothesis put forward by Paul et al. (2005) that children with autism would be likely to perform better on grammatical tasks, where options tend to the categorial, than on pragmatic/affective tasks where perceptions tend to be scalar.

The Focus Output result supports the finding of Baltaxe and Guthrie (1987) that accent on the first item is the default place for children with HFA. In unimpaired speech, default accent occurs on the final accentable syllable (Cruttenden, 1997).

Relationship Between Processing Levels and Modes

Expectations were that input and output ability would correlate, and this is broadly speaking the case (Table 3). For the typically developing group, lack of correlation in one function (Focus) may be explained by task design: although PEPS-C tasks for receptive and expressive modes mirror each other for the most part, this is least true for the Focus tasks.

In the form tasks, where correct perception of the stimulus was necessary for precise imitation, correlations were high, as predicted, for the typically developing group, but not significant in the HFA group, where children scoring at chance on discrimination tasks were as likely to produce competent imitation scores as not. Imitation scores provided an instance of the interesting but so far unexplained fact that children with HFA are poor at imitation but able to repeat large parts of their favourite videos, verbatim and in the appropriate accent.

Although it is assumed that reception precedes and affects development of expression, several studies dating from the 70s and 80s found, as Cutler and Swinney put it, "an apparent anomaly, in that young children's productive [prosody] skills appear to outstrip their receptive skills" (Cutler & Swinney, 1987, Abstract). It is however possible that, in prosody, factors other than what is heard contribute to its development. In their 1987 study, Cutler and Swinney also compared the ability to understand the functionality of prosody (the communicative role of accent) with the ability to discriminate between the relevant differences of form (variation in degrees of accent) in a small number of children, and found that the former may precede the latter. This order of skill-acquisition was also apparently the case in a small percentage of children in the current study, in both groups, and it implies that factors other than auditory processing, e.g. feedback from the child's own attempts to express the required distinctions, are active in the acquisition of functional prosodic skills. Nevertheless, Table 3 shows high correlation between receptive and expressive scores in two functions (Turnend and Chunking) and correlation approaching significance on the other two (Affect and Focus) in the HFA group. Receptive prosody skills may therefore both be acquired, although later than expected, and have an effect on associated expressive skills. Improvement in discrimination skills may also be important for receptive function skills: the relationship between discrimination tasks and function tasks suggest that Intonation Input correlated in different degrees with most receptive function scores (with Turnend Input, $r = .639$, $p < .001$; with Affect Input, $r = .454$, $p = .010$; with Chunking Input, $r = .479$, $p = .006$; with Focus Input, $r = .227$, $p < .228$). Correlations of Prosody Input with receptive function scores do not achieve significance (Turnend Input, $r = .417$, $p = .020$; Affect Input, $r = .255$, $p = .165$; Chunking Input, $r = .330$, $p = .069$; Focus Input, $r = .262$, $p < .161$).

5. Conclusion

Even with a relatively well defined experimental group, prosodic ability as measured varied considerably across individuals. All individuals in the HFA group had difficulty with at least one aspect of prosody, however, and in summary, the typically developing children's scores were significantly better than those in the HFA group. On the whole receptive and expressive prosodic skills were correlated, suggesting that targeting receptive skills in intervention may have a beneficial effect on expressive skills. Correlation of prosody scores with chronological and verbal mental age in the HFA group suggested a pattern of delay rather than deviance, although patterns of errors in some tasks suggested the latter.

One limitation of the study was the lack of acoustic analysis of atypical expressive prosody. This might clarify its relationship with poor discriminatory ability, and would help to define the nature of atypical expressive prosody in autism to a greater extent than is possible with tasks with only two or three degrees of freedom, such as those provided by the PEPS-C. This in turn would provide speech and language therapists with more focused targets for intervention.

Many of the study's findings support those of previous research, such as that accent placement is disordered in children with autism (Baltaxe & Guthrie, 1987; Shriberg et al., 2001). Findings with respect to the understanding and expression of affect were also in line with expectations (as indicated in Rutherford et al., 2002) concerning children with autism, i.e. the children with HFA had marked difficulty with these tasks whereas typically developing children did not. Results with regard to prosodic phrasing and turnend type were less striking but nevertheless suggest some difficulties for children with HFA. The finding that receptive prosodic skills may develop late in children with autism suggests that prosodic awareness could be tapped and developed in children with autism, thus possibly leading to better social skills as well as improved communication.

Acknowledgments

This research was funded by the Scottish Executive Chief Scientist Office (CZB/4/3/4). We thank the parents, speech and language therapists, teachers and children who have supported and participated in this research.

References

American Psychiatric Society (1994). *Diagnostic and statistical manual for mental disorders - fourth edition*. Washington, D.C: American Psychiatric Publishing, Inc.

Baltaxe, C.A.M., & Simmons, J.Q. (1985). Prosodic development in normal and autistic children. In E. Schopler & G.B Mesibov (Eds.), *Communication problems in autism.* (pp. 95-125). New York: Plenum Press.

Baltaxe, C.A.M., & Guthrie, D. (1987). The use of primary sentence stress by normal, aphasic and autistic children. *Journal of Autism and Developmental Disorders, 17,* 255-271.

Banse, R., & Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*, 614-636.

Baron-Cohen, S., Leslie, A.M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition, 21,* 37–46.

Baron-Cohen, S., & 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.

Boucher, J., Lewis, V. and Collis, G. M. (1998). Familiar face and voice matching and recognition in children with autism. *Journal of Child Psychology and Psychiatry, 39*, 171–181.

Butcher, A. (1981). Aspects of the speech pause: Phonetic correlates and communicative function, *Arbeitsberichte, 15.* Institut für Phonetik, Universität Kiel.

Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language, 12,* 643-661.

Cruttenden, A. (1997). *Intonation - second edition*. Cambridge: CUP.

Cutler, A., & Swinney, D. (1987). Prosody and the development of comprehension. *Journal of Child Language, 14,* 145-167.

Cutler, A. (1980) Errors of stress and intonation. In V.A. Fromkin (Ed.) *Errors in linguistic performance. Slips of the tongue, ear, pen, and hand.* San Francisco : Academic Press.

DiLalla, D.L., & Rogers, S.J. (1994). Domains of the Childhood Autism Rating Scale: Relevance for diagnosis and treatment. *Journal of Autism and Developmental Disorders, 24,* 115-128.

Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale – second edition*. Windsor: NFER-Nelson.

Fay, W., & Schuler, A.L. (1980). *Emerging language in autistic children.* London: Edward Arnold.

Fine, J., Bartolucci, G., Ginsberg, G., & Szatmari, P. (1991). The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 32,* 771-782.

Fosnot, S. M., & Jun, S. (1999). Prosodic characteristics in children with stuttering or autism during reading and imitation. In J.J. Ohala & Y. Hasegawa (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences.* (pp. 1925–1928). Dordrecht: Foris

Fry, D.B. (1958). Experiments in the perception of stress, *Language and Speech, 1,* 126–152.

Gilliam, J.E. (1995). *Gilliam autism rating scale (GARS)*. Austin, Tx: PRO-ED.

Happe, F. (1997) Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, 15, 1-12.

Katz, W. F., Beach, C. M., Jenouri, K., & Verma, S. (1996). Duration and fundamental frequency correlates of phrase boundaries in productions by children and adults. *Journal of Acoustical Society of America*, *99,* 3179-3191.

Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Leventhal, B.L., DiLavore, P.C. et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205-223.

Macintosh, K.E., & Dissayanake, C. (2004). Annotation: The similarities and differences between autistic disorder and Asperger's syndrome: A review of the empirical evidence. *Journal of Child Psychology and Psychiatry, 45, 421-434.*

McCann, J., & Peppé, S. (2003). Prosody in autistic spectrum disorders: A critical review. *International Journal of Language and Communication Disorders, 38,* 325-350.

McCann, J., Peppé, S., Gibbon, F., O'Hare, A., & Rutherford, M. (in preparation). Prosody and its relationship to language in school-aged children with high-functioning autism.

McLoone, P. (2000). *Carstairs scores for Scottish postcode sectors from the 1991 Census*. Glasgow: Public Health Research Unit, University of Glasgow.

Mozziconacci, S.J.L. (1998). *Speech variability and emotion: Production and perception*. Ph.D. thesis, Eindhoven, The Netherlands.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. (2005). Perception and production of prosody by speakers with autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 35,* 205-220.

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics and Phonetics, 17,* 345-354.

Pierrehumbert, J.B. (1987). *The phonology and phonetics of English intonation.* Bloomington, Ind: Indiana University Linguistic Club.

Roach, P. (2000). *English phonetics and phonology.* Cambridge: CUP.

Rutherford, M.D., Baron-Cohen, S., & 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.

Scott, D. R. (1982). Duration as a cue to the perception of a phrase boundary. *Journal of Acoustical Society of America,* 71, 996-1007.

Shriberg, L.D., Paul, R., McSweeny, J.L., Klin, A., Cohen, D.J., & 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., & 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.

Wells, B., & Peppé, S. (2003). Intonation abilities of children with speech and language impairment. *Journal of Speech, Language, and Hearing Research, 46,* 5-20.

Wells, B., Peppé, S., & Goulandris, A. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31,* 749-778.

Wing, L., & 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.

World Health Organisation (1993). *The ICD-10 classification of mental and behavioural disorders.* Geneva: World Health Organisation.

Appendix I
PEPS-C task descriptions and instructions for administration and scoring

<u>Intonation Input and Prosody Input: Auditory discrimination of prosodic forms without reference to meaning; on single words and short phrases.</u>
The Same-Different protocol is run before the test begins: two red circles (the Same-symbol) appear and it is agreed that the circles are the same as each other. Then the Different-symbol (a red circle and a green square) appears and it is agreed that these symbols are different.
Task: **"**We are going to listen to some noises and decide if they sound the same or different. These ones are the same; listen:" [click for sound: one laryngograph recording, repeated); same-symbol appears on left of screen and different-symbol on right] "So you have to click on Same" [tester models response by clicking on screen]. "But these ones are different; listen:" [click for sound: two different laryngograph recordings] "So you have to click on Different." Computer records child's choice and tests for match between response-choice and stimulus.
Scores: Match between stimulus and response = 1; non-match = 0.

<u>Intonation Output and Prosody output: Imitation of single words and short phrases.</u>
"You are going to hear some words from the computer. You need to try and copy the words making them sound *exactly* the same way as the computer says them. I will do the first one; listen:" [click for sound, model response by imitating computer). "So the computer speaks and then you speak. Now you try." Clients should be encouraged to listen carefully before imitating the stimulus, and to pay little attention to regional accent differences in the stimuli. The tester has a keypad with custom-designed keys, and should press G for a good response, i.e. a perfect echo of the intonation of the stimulus; F for a fair response which was not an exact imitation but maintained the function: e.g. the stimuli was a question and the client used questioning intonation, but not very clearly; and P for a poor or incorrect imitation, e.g. the stimulus was a question but the imitation was more like a statement.
Segmental or lexical errors attract no penalty.
Scores: G = 1, F = 0.5, P = 0.

<u>Turnend Input: Understanding whether an utterance requires an answer or not</u>
"You will hear the person on the computer say the names of some food. Sometimes she sounds as if she's asking if you want some food, like this:" [click for sound: "carrots" said with questioning tone. Simultaneously a response-choice screen appears: the left half shows a person holding out carrots on a plate, the right half shows a person reading from book, with carrots in a call-out] "and you click on this picture". [Tester demonstrates clicking on appropriate half]. "But if she sounds as if she is just telling you what she sees in the book, like this" [click for sound] "then you click on this picture."
Scores: match = 1, non-match = 0.

<u>Turnend Output: Producing words with intonation to suggest questioning or stating.</u>
You'll see pictures one by one on the screen. If the picture shows somebody offering some food and a question mark, like this:" [tester indicates picture, which fills screen] "you say the food as though you were asking me if I want some." [Tester models the response and presses **?** on the keypad; the next picture appears.] "If it shows someone looking at a picture of the food in a book, like this" [indicates screen] "say the food as though you were just telling me what it is." On the keypad, the tester can press ? for question, ✓ for a statement or A if the response seemed ambiguous. Items are in randomised order.
Scores: match between tester's judgment and stimulus picture = 1, mismatch/ambiguous = 0.

<u>Affect Input: Comprehending liking or disliking.</u>
"We are going to find out what kind of food the person on the computer likes. This picture shows tea" [screen shows a cup of tea and a teapot] "and she likes tea, so she says it like this" [click for sound: tea said in a happy positive way, with rise-fall tone] After the sound has played, the screen shows a happy face and a sad face. "So you click on the happy face"

[demonstrates]. A picture of mushrooms appears. "These are mushrooms; she doesn't like mushrooms so she says it like this" [click for sound: "mushrooms" said in an unhappy negative way, with fall-rise tone]. The happy and sad faces reappear. "So you click on the sad face. Now you try."
Scores: match = 1; non-match = 0.

<u>Affect Output: Producing affective intonation to suggest either liking or disliking.</u>
Tester: "I'm going to try and guess what kinds of food you like by the way you say them. Pictures of food will appear on the screen one by one. If you like the food, say the word as if you really like it, and then click on the smiley face. If you don't like it, say the word as though you don't really like that food, and then click the sad face." [Tester looks away from screen and child's face while making a judgment to avoid being influenced by non-prosodic factors.] The faces reappear. Tester: "Now click on the face that best matches how you feel about that food." This provides non-prosodic verification of the child's likes and dislikes, and is matched by the computer with the tester's judgment. Items are in randomised order.
Scores: match = 1; ambiguous judgment or mismatch = 0.

<u>Chunking Input: Comprehending prosodic phrase boundaries.</u>
"We are going to hear the computer say some sentences or phrases. You need to decide if it best matches the pictures on this side" (indicate left side of screen) "or this side" (indicate right side of screen). "On this side," (indicate left) "we have pink-and-black and green socks. On this side," (indicate right) "we have pink and black-and-green socks. Listen:" [click for sound]. "This time we have some food*:* on this side," (indicate left) "we have fish-fingers and fruit, and on this side" (indicate right) "there's fish, fingers and fruit. Listen:"
Scores: match = 1; non-match = 0.

<u>Chunking Output: Producing prosodic phrase boundaries</u>
Tester: "Pictures will appear on the screen and I want you to tell me what you see." The tester sits at right angles to the child to avoid seeing the stimuli, and presses **1** on the key pad if the first prosodic phrase-break comes after the first word of the response and **2** if it comes after the second word (food-items) or third word (colour-items).
Scores: match = 1; ambiguous judgment or mismatch = 0.

<u>Focus Input: Comprehension of contrastive accent.</u>
The screen shows one black and one blue patch of colour. Tester **"**Listen carefully. Earlier on today, the person on the computer bought some socks. But when she got home, she realised she had forgotten to buy one colour. If she says this" [click for sound]: "I wanted BLUE and black socks"] "that means she forgot to buy the blue ones, because she said 'I wanted BLUE and black socks', so you click on blue." (child click*s* on blue patch of colour). "But if she says this:" [click for sound: "I wanted blue and BLACK socks"] "that means she forgot to buy the black ones, because she said 'I wanted blue and BLACK socks', so you click on black. Items are in randomised order.
Scores: match = 1; non-match = 0.

<u>Focus Output: Production of contrastive accent.</u>

A screen appears showing animals with a football. Tester: "The cows and the sheep are playing football. This is the sheep team; they are all different colours. There is a black sheep, a blue sheep, a green sheep, a red sheep and a white sheep. And this is the cow team: there is a black cow, a blue cow, a green cow, a red cow and a white cow. There is a commentator for this football match. He tells you what is happening during the game. But he is a bit silly and gets things wrong." [Screen appears showing one green cow with football] "He is going to tell you who he thinks has the ball*,* and you have to correct him.*"* [click for sound: 'The green sheep has the ball'] You say 'No, the green COW has the ball'. Tester presses **1** if the client used contrastive stress on the colour, **2** if stress was on the animal, and **A** if response

ambiguous, i.e. hard to tell where stress was placed or if it was on a word other than the colour or animal. Items are in randomised order.
Scores: match = 1, ambiguous judgment or mismatch = 0.

Appendix II
PEPS-C task-items

<u>Turn-end Input and Output</u>: . = reading (statement); ? = offering (question).
Examples: 1.carrots? 2.tea. Practice items: 1.milk. 2.salad?
Test items: 1.cheese. 2.apple? 3.orange. 4.cake? 5.honey? 6.lemon. 7.pear. 8.jam? 9.cabbage. 10.leeks? 11.water? 12.tea. 13.carrots? 14.cream? 15.milk. 16.raisins.

<u>Affect Input and Output</u> (^ = likes; ~ = dislikes)
Examples: 1.tea^ 2.mushrooms~ Practice items: 1.milk~ 2.cream^
Test items: 1.bananas^ 2.pears^ 3.eggs~ 4.oranges~ 5.bread^ 6.jam~ 7.apples^ 8.cake^ 9.honey~ 10.raisins~ 11.salad~ 12.leeks^ 13.cabbage^ 14.carrots~ 15.cheese^ 16.tomatoes~

<u>Chunking Input and Output</u>
Examples: 1. pink&black and green socks 2.fish, fingers and fruit
Practice items: 1. black and pink&red socks 2.fish-fingers and fruit
Test items: 1.black&pink and red socks 2.cream-buns and chocolate 3.pink and black&green socks 4.black and green&pink socks 5.chocolate, biscuits and jam 6. pink&black and green socks 7.black&green and pink socks 8.chocolate-cake and buns 9.fruit, salad and milk 10.black and pink&green socks 11.chocolate, cake and buns 12.fruit-salad and milk 13.chocolate-biscuits and jam 14.red and pink&black socks 15.cream, buns and chocolate 16.red&pink and black socks

<u>Focus Input</u>: I wanted…
Examples: 1.blue and BLACK socks 2. BLUE and black socks
Practice: 1. BLUE and white socks 2. blue and WHITE socks
Test items: 1.BLACK and red socks 2.blue and GREEN socks 3.red and BLACK socks 4.BLUE and red socks 5.GREEN and blue socks 6.green and BLACK socks 7.blue and RED socks 8.GREEN and black socks 9.RED and blue socks 10.BLACK and green socks 11.green and BLUE socks 12.RED and black socks 13.blue and GREEN socks 14.BLUE and green socks 15.red and BLUE socks 16.black and RED socks

<u>Focus output</u>
Examples: 1.Now the green sheep has the ball: No, the green COW has it.
2. And the red sheep has the ball: No, the GREEN sheep has it
Practice: 1.black cow: GREEN cow. 2.green cow: green SHEEP.
Test items: 1.red sheep: red COW. 2.green sheep: RED sheep. 3.black cow: RED cow. 4.red cow: red SHEEP 5.blue sheep: blue COW. 6.red sheep: BLUE sheep. 7.black cow: BLUE cow. 8.blue cow: blue SHEEP. 9.white sheep: white COW. 10.red sheep: WHITE sheep. 11.black cow: WHITE cow. 12.white cow: white SHEEP. 13.black sheep: black COW. 14.red sheep: BLACK sheep. 15.white cow: BLACK cow. 16.black cow: black SHEEP.