Running Head: PROSODY IN CHILDREN WITH AUTISM

Receptive and Expressive Prosodic Ability in Children with High-Functioning Autism
Susan Peppé, Joanne McCann, and Fiona Gibbon
Queen Margaret University College, Edinburgh, UK
Anne O’Hare
Royal Hospital for Sick Children, Edinburgh, and University of Edinburgh, UK
Marion Rutherford
Royal Hospital for Sick Children, Edinburgh, UK

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

Address correspondence to:
Susan Peppé PhD
Speech and Hearing Sciences
Queen Margaret University College
Clerwood Terrace
Edinburgh EH12 8TS
UK
Tel: (0131) 317 3544
Email: speppe@qmuc.ac.uk

Abstract

Purpose: This study aimed to identify the nature and extent of receptive and expressive prosodic deficits in children with high-functioning autism (HFA).
Method: 31 children with HFA, 72 typically developing controls matched on verbal mental age and 33 adults with normal speech completed the prosody assessment procedure, Profiling Elements of Prosodic Systems in Children (PEPS-C).
Results: Children with HFA performed significantly less well than controls on eleven out of twelve prosody tasks ($p < .005$). Receptive prosodic skills showed a strong correlation ($p < .01$) with verbal mental age in both groups, and to a lesser extent to expressive prosodic skills. Receptive prosodic scores also correlated with expressive prosody scores, particularly in grammatical prosodic functions. Prosodic development in the HFA group appeared to be delayed in many aspects of prosody and deviant in some. Adults showed near-ceiling scores in all tasks.
Conclusions: The study demonstrates that receptive and expressive prosodic skills are closely associated in HFA. Receptive prosodic skills would be an appropriate focus for clinical intervention and further investigation of prosody and the relationship between prosody and social skills is warranted.
Receptive and Expressive Prosodic Ability in Children with High-Functioning Autism

Introduction
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 some aspect of autism may give rise to at least some aspects of prosodic disorder. By establishing whether there is an association, and, if so, the nature of the association, it may be possible to increase understanding of the 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 receptive skills have received little attention; 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 syndrome. However, the distinction between HFA and Asperger syndrome is controversial (Macintosh & Dissayanake, 2004). The most recent diagnostic criteria hold that individuals with Asperger syndrome, unlike those with HFA, do not demonstrate a general language delay during the preschool years, 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). In this study we limit our investigation to individuals with HFA but not Asperger syndrome, 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 autism and abnormal prosodic ability may be particularly closely associated. 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 (as described in the diagnostic manual DSM-IV, American Psychiatric Society, 1994), perhaps because they have difficulty with global processing (weak central coherence, Frith & Happé, 1994); this may result in a failure to access some of the non-literal aspects of language that enable social interaction. Prosody, in the sense of supralexical prosody, i.e., at utterance rather than at word level, 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 supralexical prosodic errors (e.g., pitch accent placement), although lexical stress errors often are corrected (Cutler, 1983); this suggests that frequently prosody does not impinge on speakers as being something they can or should correct. Furthermore, it appears that misleading prosody generally goes uncorrected as language develops, whereas caregivers often correct a child’s erroneous pronunciation, grammar or lexical choice (Tarplee, 1996).

A skill such as prosody which unobtrusively conveys emotional and pragmatic aspects of speech may be particularly vulnerable in children with autism, but those who do
not have learning difficulties may be capable of increasing their prosodic awareness and ability. This in turn may enhance their understanding of the subtler aspects of communication, such as the conveying of emotions, attitudes and social cues. Improvement in receptive prosodic ability may also cause an improvement in prosodic expressiveness.

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 with opposite meanings, such as monotonous and exaggerated (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. We begin by describing the terminology relevant to the measurement and interpretation of prosodic ability; 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 (e.g., Cruttenden, 1997), but what is included in the domain of prosody, and the ways in which prosodic terminology is used, can vary greatly. Because of this, the operational definition of prosody in autism tends to vary from one study to another.

The term prosody is used to describe suprasegmental characteristics of speech deriving from variations in the duration, amplitude and fundamental frequency of speech-sounds, which provide the acoustic realisations of the communicative functions conveyed by prosody. These functions have been categorised (see e.g., 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 (Profiling Elements of Prosodic Systems in Children, PEPS-C, see Appendix) 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 including pause after the phrase (Butcher, 1981); lengthening of the final syllable (Scott, 1982); and the inclusion of nuclear tone (Crystal, 1969), located at or near the end of the utterance. 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 versus chocolate-cake and cream; Ellen, the dentist is here versus Ellen, the dentist, is here) and the presence or absence of prosodic boundary markers within the utterances can resolve the ambiguities. In the PEPS-C, this demarcation function is designated Chunking.

The focus of utterances can be changed for pragmatic purposes, such as a desire for particular or contrastive emphasis, and focus is carried by accent on individual syllables. Accented syllables are relatively prominent; the acoustic features of accent are changes in fundamental frequency (such as boosted pitch) and extra length and loudness (Fry, 1958). Similar features also indicate stress, a term generally used to designate the main syllable of a word; the accent or nuclear tone that indicates completeness of an utterance is sometimes known as sentential stress. Focus is indicated by the accented word or syllable in an utterance, usually the stressed syllable of the final word; 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. In broad focus, accent on the final word or syllable may not be
noticeable, but its prominence would be increased if particular emphasis or contrast were required, in which case the focus of the utterance is described as narrow. An example is the utterance *I wanted some socks* where *socks* is heavily accented in order to contrast with a previously mentioned item such as shoes. Narrow focus can be located at any point within an utterance. 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 contentful response is required, while a falling intonation usually suggests finality (Cruttenden, 1997) requiring concurrence or acknowledgement. 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.

**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 and stress (both placement and realisation) were found to be disordered in individuals with autism in several studies (Shriberg, Paul, McSweeney, 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, 1985). McCann and Peppé (2003) suggested two main reasons for these contradictions. First, diagnosis of autism was not consistent for all study participants. Secondly, there was 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.

Some relevant aspects of prosody (e.g., receptive abilities) were seriously under researched in the literature. Given the unusual quality of expressive prosody in autism, it is not surprising that researchers have directed their attention towards prosodic expression, but it would 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.

A more recent paper (Paul, Augustyn, Klin, & Volkmar, 2005) investigated perception and production of prosody in children with autism and this study is considered here 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. The participants in the study by Paul et al. were older than those in the current study (mean chronological age of 16+ years). Results suggest little difference between the performances of the participants with autism and the controls on most of the tasks, but the authors concluded that this was less because of a lack of difference in prosodic ability and more because of weaknesses of task design. One problem was ceiling effects, which were apparent in both groups on some tasks.
Where ceiling effects occurred in perception tasks, the authors suggested that prosodic differences in the task items may have been too apparent, and 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 pointed 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. Other reasons for lack of group differences could have been that the age and clinical diagnosis of the participants was wide ranging, or 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 prominence, whether pragmatically (at the utterance level, for purposes of contrast) or for grammatico-lexical distinction (lexical stress), 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 methodology is taken to be more responsible than a mistaken premise for lack of substantiation. The present study aimed to avoid as many as possible of the pitfalls of previous studies by recruiting participants in the experimental group with a narrow diagnosis of autism; by having adequate numbers of both experimental and control participants; and by investigating a broad range of prosodic ability using tasks of similar design, thus enabling abilities in all aspects to be compared with each other. The hypothesis remains the same as in previous studies; that prosodic ability is lower in children with HFA than in typically-developing children. With intervention in mind, another aim was to look for correlations between receptive and expressive abilities in individual aspects of prosody, and to analyse between-group differences in terms of deviance (where error-patterns differ) or delay (where the prosodic performance of children with HFA resembled that of younger typically-developing children), these being beneficial as an indication of clinical priorities.

**Method**

**Participants**

Participant groups were as follows: 31 children with HFA, 24 boys, 7 girls, age range 6;1 to 13;6 (M = 9;10, SD = 2.3); 72 children with typical development, 54 boys, 18 girls, age range 4;10 to 11;8 (M = 6;10, SD =1.5); and 33 adults, 10 male, 23 female, age range 18-59 (M = age group 30-39). Criteria for inclusion in the group with HFA were as follows: a diagnosis of autism (but not Asperger syndrome) in preschool years; current chronological 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 group with TD were that they should have a verbal mental age > 4;0 years and socioeconomic status similar to that of the children with HFA. To achieve the latter, the study used the Carstairs Index (McLoone, 2000) as a classification of
Prosody in children with autism

deprivation. This index ensured that the typically developing children were drawn from Edinburgh state primary schools that were in the same postcode areas as the children with autism. The group with TD 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. A total of 76 controls were recruited to ensure a likelihood that at least one child would be matched with each child with HFA on both verbal and socioeconomic measures. (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.) Out of 76 children, 95% met the verbal mental age criteria and were included in the study.

Adults who completed the PEPS-C test had normal speech and language and met all criteria that applied to the typically developing children, except that some were resident in regions of Scotland other than just Edinburgh. In these areas, however, the intonation of the regional varieties was similar to that of Edinburgh.

**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 syndrome were not selected.

**Procedures**

Findings reported in this study for all children 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 and tests of pragmatic and non-verbal ability as part of a larger research project; results of these are reported in a separate paper (McCann, Peppé, Gibbon, O’Hare, & Rutherford, submitted).

**Prosody Assessment**

PEPS-C is a relatively new prosody assessment procedure, although previous studies employed a manual version (Wells & Peppé, 2003; Wells, Peppé, & Goulandris, 2004). The computerized Scottish version is described in Peppé and McCann (2003) and details of tasks, instructions, scoring procedures and task items are in the Appendix. The procedure aims to evaluate prosodic skills according to a psycholinguistic model (Stackhouse & Wells, 1997). Tasks are at two levels: that of form processing, requiring non-cognitive skills (auditory discrimination and the production of prosodic variation through imitation), and that of communicative function (pragmatic or interactional, affective, and grammatical or linguistic), involving cognitive processes. Tasks are also in two communicative modes (i.e., direction of communication): receptive and expressive.
Function Tasks

In parallel reception and expression 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). The test elicits all responses by pictures and uses items that are culturally inoffensive. The pictures included in the test are likely to be familiar to young children, and before the test the children completed a vocabulary check to ensure that they could name the pictures. Segmentally, the items were easy to pronounce. For the reception 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 reception 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 50%. In expression tasks there was also a possibility of producing the right answer by chance, so the competence level was similarly set at 75%.

Form Tasks

These were divided into Short-item tasks (one or two syllables) and Long-item tasks (six or seven syllables)*. Discrimination tasks follow a Same-Different format, and stimuli are from laryngograph signals taken from a microphone placed on the speaker’s throat and recorded simultaneously with the function task stimuli. For the Short-item Discrimination task, equal numbers of stimuli from the Affect and Turnend function tasks were used; for the Long-item Discrimination task, equal numbers of Chunking and Focus stimuli. The purpose of these tasks was to discover whether the children’s auditory processing was at a level that would enable them to perceive the acoustic differences used in the function reception tasks. Expression (Imitation) tasks sought to elicit a prosodic repertoire by requiring participants to imitate stimuli that were similar to stimuli used in the reception function tasks, but with some additional aspects (e.g., more tones, more varied stress-placement).

Assessment Conditions

A certified speech and language therapist (the second author) interviewed the children with HFA in a clinic room designed for recording speech and language therapy sessions; parents and siblings could be present if wished. Interviews with typically developing children mainly took place in a quiet room in their schools. Both groups were interviewed by the same person using the same equipment. The interviewer explained each prosody task and administered two practice items per task. If the children failed the practice tasks they were re-administered and if these failed, the task was abandoned. Sessions were recorded on DAT tape and CD with computer backup. Administration of the PEPS-C tasks took between 40 minutes and an hour, depending on the child’s age, ability and attention span. Children with HFA took two one-hour sessions to complete the battery of assessments. Children sat approximately 18 inches from the computer screen, and were free to leave the assessment whenever they wished. Caregivers accompanied the children with HFA and could say if necessary whether the child wished to cease participation. As the children with autism and

* These tasks are designated Intonation and Prosody respectively in the beta-version of PEPS-C currently in use by some researchers.
the typical children completed a different battery of tests it was not practical for the tester to
be blind to the diagnosis.

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.
Examples of what is seen in the pictures are given in the Appendix. 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 to
them. After each 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 some circumstances, the tester replayed the item, for example, where children
had been distracted, or had not heard the item due to ambient noise, or if children said they
had made a wrong mouse click. The Appendix gives details of instructions and scoring
procedures, and a demonstration of the program, with examples of the pictures used, is
located on the website: http://www.qmuc.ac.uk/ssrc/prosodyinasd/.

Statistical Methods

Analysis of covariance established significance of group differences on task scores
and 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. Partial eta squared indicated effect size. Pearson’s correlations established the
relationships between the measures used. Post-hoc analyses (t-tests) identified differences in
error types as significant if p < .01.

Reliability

Only the expression scores were subject to rater variation. For a measure of intrarater
reliability, the tester re-judged 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. Re-judged data consisted of the complete set of
expression 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 (Long-item Imitation). We used the earlier set
of task scores to calculate the children’s performance and inter-rater comparisons.

As an inter-rater 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), maximum .98 (Short-item Imitation).

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. The groups did not differ significantly on either social deprivation category or sex.

**Prosody Task Scores**

We compared the mean percentage prosody scores of the HFA group with those of the TD group; results are shown in Table 1, together with adult scores. The HFA group’s mean scores are significantly lower than those of the TD group in seven of the twelve tasks: Affect Reception, Affect Expression, Focus Expression, and the four form tasks (Short-item Discrimination, Short-item Imitation, Long-item Discrimination, and Long-item Imitation). The TD group has mean scores above competence level, but without ceiling effects, in the same seven tasks; mean scores were approaching competence on two expression tasks (Chunking and Turnend) and below 70% on the remaining three reception tasks (Chunking, Focus, and Turnend). The HFA group’s mean scores were approaching competence on one task (Affect Reception), and below 70% on the other eleven. It is worth noting that the participants are matched on verbal mental age; if, as in normal clinical assessment, they were matched on chronological age, the differences would be more pronounced. The ranges of scores demonstrate that individuals’ performance, particularly in the HFA group, varied widely.

**TABLE 1 ABOUT HERE**

**Practice Effects**

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 (Short-item Imitation, $t(72) = 3.6, p = .001$; Long-item Discrimination, $t(72) = 2.99, p = .004$).

**Acquisition of Prosody and Language Skills**

Table 1 also shows the mean scores of 28 adults, which indicate that children may reasonably be expected to acquire these skills eventually. Table 2 shows the degree to which prosody skills correlate with chronological age in typically developing Edinburgh children and in the HFA group. It is noticeable that prosody correlates much better with chronological age in the TD group. Non-verbal ability scores correlated at the .01 level ($r = .502$) with receptive prosody scores but not significantly with expressive scores.

**TABLE 2 ABOUT HERE**

**Deviance in Development of Prosodic Ability**

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

**Form Tasks**

In the Short-item Discrimination 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). In Long-item Discrimination tasks, this difference was even greater ($t = 6.52, p < .001$).

**Turnend.** In the reception task, there was no significant difference in judgment preferences between groups, but some 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). In the expression 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 ($t = 2.34, p = .019$). For example, *carrots* was likely to be uttered with sharply rising intonation in the HFA group, even in response to the picture of a child reading the word.

**Affect.** In the reception task, the HFA group made more errors of both types than the TD group: in failing to discern liking a food item (see Appendix 1) ($t = 3.97, p < .001$), and in failing to discern dislike of it ($t = 2.93, p = .003$). In the expression 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 significantly more frequent than for those in the TD group, both for liking ($t = 3.57, p < .001$) and for disliking ($t = 2.33, p = .02$). In other words, expression of affect was fairly inscrutable in the HFA group.

**Chunking.** In the reception task there were no discernible error patterns. In the expression 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. They were therefore more likely to sound as though they were saying *fruit-salad and milk* rather than *fruit, salad and milk*.

**Focus.** In the reception task, the HFA group made significantly more errors than the TD group when the accent was early, i.e., on the colour (see Appendix) than when it was later (on the animal) ($t = 4.36, p < .001$). In the expression task, their responses showed a significantly greater tendency than the TD group to place accent on the colour ($t = 3.5, p < .001$), i.e., in response to the commentator’s *The red cow’s got the ball*, the children with HFA would say *No, the red sheep’s got it* (emphasized word in bold). HFA responses of both types were significantly more ambiguous than those of the TD group (accent on colour: $t = 3.43, p = .001$; accent on animal: $t = 5.2, p < .001$).

**Receptive and Expressive Scores**

Correlations between receptive and expressive scores in the HFA and TD groups are shown in Table 3. 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, correlation was weak or not significant between reception and expression on tasks involving polysyllabic utterances: Chunking ($r = .095$) and Focus ($r = .260$); but strong for tasks involving shorter items: 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 or affective functions (Affect ($r = .382, p = .019$) and Focus ($r = .317, p = .047$). Within form tasks, findings for the HFA group showed that discrimination (reception ability) did not correlate with the ability to imitate utterances (expression ability); for the TD group, by contrast, correlation in form tasks was highly significant (Intonation $r = .459, p < .001$; Prosody $r = .557, p < .001$).

**Discussion**

For the HFA group, overall mean prosody scores correlated only weakly with chronological age ($r = .378, p = .036$, Table 2). From receptive and expressive overall mean scores, it can be seen that the bulk of the correlation is carried by the receptive scores ($r = .597, p < .001$). Receptive mean scores also correlated strongly with verbal mental age ($r = .775, p < .001$, Table 2) and, as already noted, with non-verbal ability. Because the measure
Prosody in children with autism

of VMA was a receptive task, it might be expected to correlate less well with expressive than with receptive prosody skills, and this is the case (Table 2). This suggests that maturation alone may result in improved receptive prosody, especially where non-verbal ability is good, but that understanding of prosody is strongly linked with verbal ability (as exemplified by understanding of lexis), bearing in mind that if this link is causal we have not established which way the causation acts. The correlation between PEPS-C receptive task scores and chronological and verbal mental age suggests 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 of these children in this respect would appear to be deviant compared with that of typically-developing children: Table 2 shows strong correlations between chronological and verbal mental age in the TD group.

**Auditory Processing in Form Tasks**

Table 1 showed that children with autism performed significantly less well than typically developing peers on reception form tasks (auditory discrimination), and that their tendency was to perceive difference in Same items. The opposite error (perceiving Different items as Same) would, more simply, suggest a lack of ability to perceive auditory differences, but the error found in these results is considered deviant and requires investigation. 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 explain 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. How this ability may differ from typically-developing children is perhaps a subject for neuroimaging studies, but is beyond the scope of the present paper.

**Prosodic Functions**

The Turnend tasks show that children with autism had a tendency to judge questions as statements 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 reception stimuli according to their own preferences, despite reminders that they were to consider the preferences of the speaker on the computer. High ambiguity for both options in the expression 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 reception task, nor in the number of responses judged as ambiguous in the expression task. This is perhaps some evidence in favour of the hypothesis put forward by Paul et al. (2005) that children with autism would perform better on grammatical tasks, where options tend to the categorial, than on pragmatic or affective tasks where perceptions tend to be scalar.

The Focus Expression result supports the finding of Shriberg et al (2001) that accent placement is disordered in autism. It also supports that of Baltaxe and Guthrie (1987), that
Prosody in children with autism

Accent on the first item is the default place for children with HFA: in unimpaired speech, default accent occurs on the final acceptable syllable (Crystal, 1969).

Relationship Between Processing Levels and Modes

Expectations were that receptive and expressive ability would correlate, and this is broadly speaking the case for both groups (Table 3), especially in function tasks. Less significant correlation for both groups 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 (see Appendix).

In the form tasks, where correct perception of the stimulus was necessary for precise imitation, correlations were high, as predicted, for the TD group, but not significant in the HFA group, where children scoring at chance on discrimination tasks were as likely as not to produce imitation scores at competence level. This dissociation between receptive and expressive skills reflects the difficulties encountered in social situations involving people with autism, and the relationship between prosodic skills and a measure such as the Social Reciprocity Algorithm of the ADOS (Lord et al. 2000) might repay investigation. Another interesting aspect, so far unexplained, is that children with autism are frequently able to repeat large parts of their favourite videos, verbatim and in the appropriate accent, i.e., with matching prosody. These children might therefore have been expected to perform well on imitation of prosody, but instead their performance on these tasks is in line with the evidence for problems of motor imitation (Williams, Whiten, Suddendorff, & Perrett, 2001). On the other hand, this result may reflect the difference between the volitional nature of quoting favourite videos and the lack of obvious purpose in performing the PEPS-C task.

Although it is assumed that receptive ability precedes and affects development of expressive ability, 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” (Abstract, Cutler & Swinney, 1987). 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.

Our results suggest that receptive prosody skills may be acquired in children with HFA, although later than expected, and that receptive and expressive skills are related. One possibility that this presents for intervention is that receptive skills could be targeted by exercises designed to enhance sensitivity to changes in prosody and make explicit the role of prosody in communication functions. The object of this would be twofold: to increase awareness of the functions of prosody in language and hence the understanding that conversation conveys information that is not strictly linguistic and may reflect personal attitudes; and secondly, because of the receptive-expressive correlation, to hope that an increase in receptive prosodic skills might result in amelioration of expressive prosodic disorder. To see if this is at least feasible, exercises based on the PEPS-C test were trialled with two participants in the experimental group (after they had completed all the tests for the study). Although no conclusions can be drawn from such a small trial, it proved possible to
devise variations on the PEPS-C tasks and to deliver and score them with ease by computer. This approach, via receptive skills, is in line with that adopted by Baron-Cohen (2004). The alternative is to address expressive prosodic disorder directly, but this study revealed a difficulty for this approach. A considerable degree of heterogeneity was apparent; within the group, mean scores on expressive prosody tasks varied widely, and within individuals the same diversity of scores could be seen. Moreover, the manifestations of atypical expressive prosody (e.g., whether pitch patterns were monotonous or exaggerated) also varied widely. Unless programmes of intervention for expressive prosody were tailored to the individual, requiring rather more knowledge of prosody than is currently available, it seems quite possible that adverse consequences (i.e., prosody that sounds even more unusual or inappropriate) could result.

Conclusion

Even with a relatively well defined experimental group, prosodic ability varied considerably across individuals. All individuals in the HFA group had difficulty with at least one aspect of prosody, however, and the typically developing children’s scores were, in summary, significantly better than those in the HFA group. As expected, and as indicated by the findings of Rutherford, Baron-Cohen and Wheelwright (2002), the HFA group differed from the TD group most in the Affect tasks, which reflected the least linguistic function of prosody. The other function task where great difference appeared was the Focus expression task, thus supporting the findings of previous research that accent placement is disordered in children with autism (Baltaxe & Guthrie, 1987; Shriberg et al., 2001). Results with regard to prosodic phrasing and turnend type were less striking but nevertheless suggest some difficulties for children with HFA. Receptive and expressive prosodic skills were correlated in function tasks, suggesting that targeting receptive skills in intervention may have a beneficial effect on expressive skills. The dissociation between receptive and expressive skills in the form tasks for the HFA group, where scores were very dissimilar to those for the TD group, suggests a possible focus for further research into the processing of prosody and the capacity for imitating it in autism. The correlation of prosody scores with chronological and verbal mental age in the HFA group suggests 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 would help to define the nature of atypical expressive prosody in autism to a greater extent, and provide more focused targets for intervention.

The indication that receptive prosodic skills may develop late in children with autism suggests that prosodic awareness could be tapped and developed in children with autism. Given the relevance of prosody to affective, non-linguistic, aspects of communication, increased awareness of the role of prosody could possibly lead 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
Prosody in children with autism


Table 1. Results of analysis of covariance between HFA and TD groups for PEPS-C scores

<table>
<thead>
<tr>
<th>PEPS-C tasks</th>
<th>F (partial η²)</th>
<th>HFA group (N = 31)</th>
<th>TD group (N = 72)</th>
<th>Adult group (N = 33)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>df 1,97</td>
<td>mean score (SD)</td>
<td>range</td>
<td>mean score (SD)</td>
</tr>
<tr>
<td>Short-item Discrimination</td>
<td>7.01** (.07)</td>
<td>68.8 (22)</td>
<td>25.0 - 100</td>
<td>80.1 (17.4)</td>
</tr>
<tr>
<td>Short-item Imitation</td>
<td>10.94** (.10)</td>
<td>64.7 (20.6)</td>
<td>9.4 - 93.8</td>
<td>79.9 (18)</td>
</tr>
<tr>
<td>Turnend Reception</td>
<td>.01 (.00)</td>
<td>65.9 (21.4)</td>
<td>18.7 - 100</td>
<td>64.8 (18.1)</td>
</tr>
<tr>
<td>Turnend Expression</td>
<td>1.76 (.02)</td>
<td>68.1 (21.8)</td>
<td>43.7 - 100</td>
<td>74.2 (18.1)</td>
</tr>
<tr>
<td>Affect Reception</td>
<td>16.21*** (.14)</td>
<td>71.2 (21.6)</td>
<td>31.3 - 100</td>
<td>84.5 (11.4)</td>
</tr>
<tr>
<td>Affect Expression</td>
<td>11.32** (.10)</td>
<td>63.3 (26.3)</td>
<td>18.7 - 100</td>
<td>79.4 (19.2)</td>
</tr>
<tr>
<td>Long-item Discrimination</td>
<td>15.1*** (.13)</td>
<td>63.5 (23.1)</td>
<td>31.3 - 100</td>
<td>79 (13.3)</td>
</tr>
<tr>
<td>Long-item Imitation</td>
<td>36.83*** (.27)</td>
<td>65.7 (22)</td>
<td>34.4 - 100</td>
<td>85.4 (11.4)</td>
</tr>
<tr>
<td>Chunking Reception</td>
<td>.04 (.00)</td>
<td>67.5 (15.7)</td>
<td>43.7 - 93.8</td>
<td>69 (15.6)</td>
</tr>
<tr>
<td>Chunking Expression</td>
<td>1.35 (.01)</td>
<td>66.5 (26.4)</td>
<td>18.7 - 100</td>
<td>71.2 (11.8)</td>
</tr>
<tr>
<td>Focus Reception</td>
<td>2.53 (.02)</td>
<td>59.6 (19)</td>
<td>31.3 - 100</td>
<td>65.9 (19.1)</td>
</tr>
<tr>
<td>Focus Expression</td>
<td>45.8*** (.32)</td>
<td>61.6 (26.4)</td>
<td>18.7 - 100</td>
<td>84 (15)</td>
</tr>
</tbody>
</table>

a HFA: N = 30 for all Expression tasks except Focus (N = 29)
b TD: N = 71 for Focus Reception and Focus Expression

*** = significant at .001 level, ** = significant at .01 level
### Table 2. Bivariate correlations of PEPS-C scores with participants’ chronological age (CA) and verbal mental age (VMA).

<table>
<thead>
<tr>
<th>PEPS-C task</th>
<th>HFA group</th>
<th>TD group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (SD)</td>
<td>range</td>
</tr>
<tr>
<td>Overall Mean</td>
<td>65.5 (20.0)</td>
<td>9.4 - 100</td>
</tr>
<tr>
<td>Overall Reception</td>
<td>66.1 (19.8)</td>
<td>18.7 - 100</td>
</tr>
<tr>
<td>Overall Expression</td>
<td>65.0 (20.8)</td>
<td>9.4 - 100</td>
</tr>
</tbody>
</table>

** = significant at .01 level, * = significant at .05 level
Table 3. Bivariate correlations between receptive and expressive scores for PEPS-C tasks

<table>
<thead>
<tr>
<th>PEPS-C task</th>
<th>HFA</th>
<th>TD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$r$</td>
</tr>
<tr>
<td>Overall</td>
<td>0.56**</td>
<td>0.73**</td>
</tr>
<tr>
<td>Turnend</td>
<td>0.74**</td>
<td>0.60**</td>
</tr>
<tr>
<td>Affect</td>
<td>0.38</td>
<td>0.58**</td>
</tr>
<tr>
<td>Chunking</td>
<td>0.61**</td>
<td>0.09</td>
</tr>
<tr>
<td>Focus</td>
<td>0.32</td>
<td>0.26</td>
</tr>
<tr>
<td>Short-item</td>
<td>-0.00</td>
<td>0.46**</td>
</tr>
<tr>
<td>Long-item</td>
<td>0.13</td>
<td>0.56**</td>
</tr>
</tbody>
</table>

$r$ = Pearson’s correlation coefficient

** = significant at .01 level, * = significant at .05 level
Appendix. Profiling Elements of Prosodic Systems in Children (PEPS-C) tasks with examples and instructions for administration and scoring.

Before the test begins, pictures used in the test appear on the screen to ensure that participants know the names of them. The Same-Different protocol is run: two red circles (the Same-symbol) appear, and it is agreed that the circles are the same as each other. The Different-symbol (a red circle and a green square) appears: it is agreed that these symbols are different.

**Short-item Discrimination and Long-item Discrimination.** Auditory discrimination of prosodic forms without reference to meaning; on single words (short-items) and short phrases (long-items). These items involve laryngograph recordings only; examples can be heard on [http://www.qmuc.ac.uk/ssrc/prosodyinasd/](http://www.qmuc.ac.uk/ssrc/prosodyinasd/)

Tester says: “We are going to listen to some noises and decide if they sound the same or different. These ones are the same, listen” [click to hear two identical stimuli; 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 Same-symbol]. “But these sounds are different, listen:” [click to hear two different stimuli] “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.

**Short-item Imitation and Long-item Imitation:** Imitation of single words and short phrases. Short-item sound stimuli examples: 1. apple (question) 2. jam (disliking) 3. eggs (liking) 4. yogurt (reading)

Long-item sound stimuli examples: 1. green, and red&black socks (comma indicates phrase boundary; & indicates no phrase boundary between colours) 2. green and black socks (bold indicates stress) 3. cream-buns and chocolate  4. red and blue socks.

“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 (good) = 1, F (fair) = 0.5, P (poor) = 0.

**Turnend Reception: Understanding whether an utterance is a question or a statement**

Sound stimuli examples: 1.carrots? (with questioning intonation, as if offering) 2. tea. (with declarative intonation, as if reading) 3. milk. 4. salad?

“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 first sound: “carrots” said with questioning tone. Simultaneously the response choices appear: the left half of the screen shows a person offering 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 second sound] “then you click on this picture.”
Prosody in children with autism

Scores: match = 1, non-match = 0.

**Turnend Expression:** Producing words with intonation to suggest questioning or stating.


“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: tester’s judgment and stimulus picture: match = 1, mismatch or ambiguous = 0.

**Affect Reception:** Comprehending liking or disliking.

Sound stimuli examples: 1. tea^ (^ = said as though speaker likes it) 2. mushrooms~ (~ = said as though speaker dislikes it) 3. milk~ 4. cream^

“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.

**Affect Expression:** Producing affective intonation to suggest either liking or disliking.

Picture stimuli examples: 1. bananas 2. pears 3. cheese 4. jam

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 to avoid being influenced by non-prosodic factors (e.g., facial expression) and enters judgment on keypad.] 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 where likes and dislikes are easily predicted (e.g., chocolate) are not included.

Scores: match = 1; ambiguous judgment or mismatch = 0.

**Chunking Reception:** Comprehending prosodic phrase boundaries

(The colour-grouping idea was inspired by the work of Katz, Beach, Jenouri, & Verma, 1996)

Sound stimuli examples: 1. pink&black, and green socks (comma indicates phrase boundary; & indicates no phrase boundary between colours. Pictures of two pairs of socks: the first bicoloured – pink and black - the second unicolour - green.) 2. chocolate, cake, and fruit (Pictures: the three food items) 3. black, and pink&red socks. (Pictures: (two pairs of socks: the first unicolour, the second bicoloured) 4. chocolate-cake, and fruit (the two food items)

“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&black, and green socks. On this side,” (indicate right) “we have pink, and black&green socks. Listen:” [click for sound]. “This time we have some food: on this side,” (indicate left) “we have chocolate-cake, and fruit, and on this side” (indicate right) “there’s chocolate, cake, and fruit. Listen:”
Scores: match = 1; non-match = 0.

**Chunking Expression: Producing prosodic phrase boundaries**

Picture stimuli examples: 1. red and black&pink socks  2. fruit, salad and cream 3. pink&red and black socks 4. fruit-salad and cream

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 (e.g., red, fruit); 2 if after the second word in food-items (e.g., fruit-salad) or third word in sock-items (e.g., pink-and-red) and A if the place of prosodic break is ambiguous. Items are in randomised order.
Scores: match = 1; ambiguous judgment or mismatch = 0.

**Focus Reception: Comprehension of contrastive accent.**

Sound stimuli: 1. I wanted blue and **black** socks (bold indicates stressed syllable). 2. I wanted **blue** and black socks. 3. I wanted **blue** and white socks. 4. I wanted blue and **white** socks.

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 clicks 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.”
Scores: match = 1; non-match = 0.

**Focus Expression: Production of contrastive accent.**

Examples:
1. Cue: Now the green sheep has the ball. (Picture: green cow.)
   Response: No, the green **cow** has it.
2. Cue: And the red sheep has it. (Picture: green sheep.)
   Response: No, the **green** sheep has it
4. Cue: ….green cow…. Picture and response: **green** **sheep**.

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: black, blue, green, red and white. 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 which animal 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 the animal. Items are in randomised order. Scores: match = 1, ambiguous judgment or mismatch = 0.