

Functionality and Perceived Atypicality of Expressive Prosody in Children with Autism Spectrum Disorders

Sue Peppé¹, Pastora Martinez Castilla⁴, Robin Lickley¹, Ineke Mennen¹, Joanne McCann¹, Anne O'Hare^{2,3}, Marion Rutherford²

¹Queen Margaret University College, Edinburgh; ²Royal Hospital for Sick Children, Edinburgh;
³University of Edinburgh; ⁴Universidad Autonoma Madrid
speppe@qmul.ac.uk

Abstract

People with autism are perceived to have 'odd' prosody, but is it malfunctioning? A new prosody test assesses the functionality of prosody in four aspects of speech (phrasing, affect, turn-end and focus) by tasks that elicit utterances in which prosody alone conveys the meaning. The test was used with 100 typically-developing children (TD), 39 with Asperger's syndrome (AspS) and 31 with high-functioning autism (HFA). In results, HFA<TD on all six tasks, HFA<AspS on four, and AspS<TD on one. In perception experiments, judges rated the atypicality of the prosody in samples of conversation from participants in each of the three groups. Correlation between the judges' ratings was high, and ANOVAs showed differences between groups similar to those found in the test results. The ratings correlated significantly (mainly at the 0.01 level) with the test's output scores. The findings support the ecological validity of the test for use as a clinical assessment tool.

Introduction

The original description of autism included the characteristic of 'odd' or atypical prosody [1], but the terms used to describe it are vague and ill-defined [2]. There are few studies to date investigating prosodic ability in this population: a recent review [3] identified 16 between 1980 and 2002, from which the findings were inconclusive and sometimes conflicting. The question arises as to how far the atypical prosody of children with autism spectrum disorders (ASD) affects their ability to communicate, i.e. whether their prosody is linguistically misleading or ambiguous, or whether the effect is simply socially disadvantageous.

To date there have been few assessment procedures designed to investigate prosody. A recent one (PEPS-C [4]) tests receptive and expressive prosodic skills in parallel tasks, and examines four linguistically communicative functions in speech (phrasing, affect, turn-end and focus, see **Prosody test**) in which prosody has a crucial role. The procedure seeks to establish whether, in an environment of structured tasks and elicited responses targeting the four communicative functions, testees' prosody is intelligible, ambiguous or misleading, and whether they can imitate the prosody as well as the words of heard utterances. With this procedure, it is possible to gain a measure of the functionality of the testees' prosody.

The second question addressed in this study therefore relates to whether the prosody assessment procedure is ecologically valid; if so, it would have potential for use as a clinical assessment tool. To investigate this, two perception experiments were carried out. Samples of conversational utterances were presented to two sets of listener-judges and

rated for atypicality, to see how far the procedure's results reflected perceptions of atypical expressive prosody in non-elicited conversational speech.

Method

In a recent study in Edinburgh, Scotland, PEPS-C data was collected from 70 children aged 5-14 with ASD. 39 were diagnosed with Asperger's syndrome (AspS) and 31 with high-functioning autism (HFA). It is controversial as to whether these two forms of ASD can be reliably distinguished, but in general (and for the purposes of the study), children in both groups have cognitive (non-verbal) functioning within the normal range, but children who have no clinically significant pre-school language delay are identified as having AspS, whereas children who do are classed as having HFA. Children with ASD completed a test on non-verbal ability [5] and only those who fell within the normal range were included in the study. As the PEPS-C test is not yet standardised, prosody data was also collected from 100 children with typical development (TD) in the same age range as the children with ASD from schools matched for socioeconomic status and regional accent. In order to compare prosodic ability with other language factors, the children with TD were matched with the children with ASD for verbal mental age on a measure often used in studies of this kind, the BPVS-II [6].

Prosody test (PEPS-C)

As mentioned above, the communicative functions tested are phrasing, affect, turn-end and focus. Only expressive tasks are described here, but the test includes parallel receptive tasks for each expressive task. For examples of the auditory stimuli and pictures used in the tasks, see www.qmul.ac.uk/ssrc/prosodyinasd/. In function tasks, judgments are of whether the distinctions are intelligible, misleading or ambiguous. In form tasks, responses are rated good, fair or poor.

Phrasing ('chunking') function

This function refers to the prosodic phrasing associated with syntactic boundaries, in which pauses, final lengthening and the presence of accent or tone combine to indicate phrase-ends. The tasks make use of lexically ambiguous phrases that can be disambiguated by prosody, with the different meanings rendered pictorially. For example, 'chocolate cake and jam' can have a phrasal break after 'cake' and be depicted as a picture of a chocolate cake and one of jam, or it can have a break after 'chocolate' and be illustrated as separate pictures of chocolate, cake and jam. Pictures of such items appear on a computer screen; the testees say what they see. The tester, not

seeing the screen, records her opinion of which pictures the testee was seeing.

Affect function

Feelings about food items are used as an instance of affective prosodic function, with participants required to indicate whether they like or dislike certain foods. Food items appear and the testees say them in a way that indicates whether or not they like it, using their own feelings as a guide. According to the intonation used, the tester assigns the child's feelings to one of the two options, and the faces then reappear so that testees can confirm their feelings by clicking on the appropriate face. This provides independent verification of the testee's target.

Turn-end function

At conversational turn-ends, intonation signals the speaker's expectations of the listener. The PEPS-C investigates testees' ability to distinguish questions and declaratives. Testees see either a picture of someone offering some food or one of someone looking at a picture of the same food in a book and are asked to say the food item as if in that situation.

Focus function

For the PEPS-C Focus task, the function of contrastive stress is used and involves correction, with auditory stimuli such as "Now the green cow has it..." where a picture shows a white cow, thus inviting the correction: "No, the *white* cow has it", with contrastive stress on 'white'. Conversely, the cue might be: "Now the white sheep has it...", prompting the response: "No, the white *cow* has it". The expected place of accent in the minimal-pair responses is thus on either the colour or the animal.

Form tasks

Additionally, the PEPS-C assesses the ability to imitate prosody: testees are given examples of items from the function tasks and asked to imitate them, copying not only the words but also the way they are said. Two tasks are used:

imitation of the intonation of single words (one or two syllables) and of short phrases (6-7 syllables).

Perception procedure

Samples of conversation of 10 seconds duration were taken from 29 of the children with AspS, 29 of those with HFA and 25 of the children with TD. The conversational material was taken either from the telling of a story shown in pictures or from conversation around the assessment tasks. Stimuli were presented in randomised order to 5 adults with phonetic awareness and Scottish accents similar to those of the children, who used direct magnitude estimation [7] to rate samples for atypicality. They were asked how odd each sample sounded, and heard at regular intervals an utterance that acted as a modulus by which to compare the samples. Additionally, a group of 22 phonetically naïve adult judges with a minimum of 3 months' exposure to the regional accent of the children rated a subset of the samples, taken from each of 5 children with HFA, 5 with AspS and 5 with TD, each group with a mean chronological age of 9 years.

Results

PEPS-C

Reliability for the scoring of the PEPS-C was examined. As an intra-rater measure, the tester re-judged 10% of responses six months after first scoring: this showed a mean 95.5% agreement on task scores. The same 10% of tasks was scored by an experienced speech and language therapist under similar conditions as an inter-rater measure: mean agreement across task scores was at 88.6%.

Differences between the verbal mental age of the groups were not significant, but were apparent between the chronological ages (HFA>TD, $t(1,128) = 3.11, p = .005$; AS>TD $t(1,133) = 2.85, p = .002$), suggesting language delay in both ASD groups. Results on the PEPS-C tasks are shown in Figure 1.

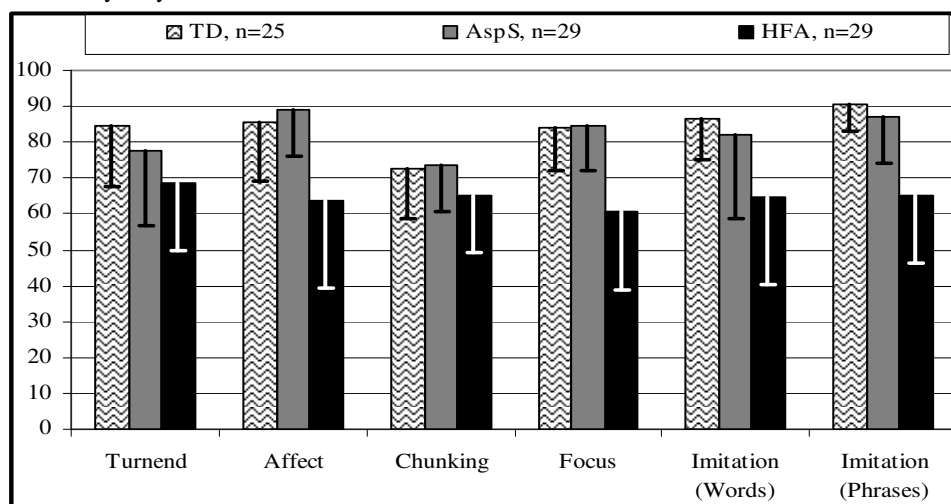


Figure 1: Mean scores on PEPS-C output tasks for each group, with error bars showing standard deviations.

- Using Tukey post-hoc tests (*HSD*), where Levene tests suggested equality of variance, the HFA group scored significantly lower than the TD group on Affect: $HSD = 3.545, p < .001$; Focus: $HSD = 4.087, p < .001$, Turn-end

($HSD = 1.903, p = .011$) and Chunking ($HSD = 1.278, p = .025$). Using Games-Howell (*G-H*) post-hoc tests, to allow for inequality of variance: imitation of words, $G-H = 3.303, p < .001$; imitation of phrases $G-H = 4.15, p < .001$.

- The HFA group also scored significantly lower than the AspS group on imitation of words: $G-H = 2.6$ ($p = .021$) and greatly lower ($p < .001$) on Affect ($HSD = 3.939$) Focus ($HSD = 3.36$) and imitation of phrases ($G-H = 3.184$).
- Difference between the TD and AspS groups was just significant on only one task (imitation of phrases): $G-H = .966$, $p = .049$.

Perception experiments

Table 1 shows results for these. In mean ratings of atypicality, there was good agreement between the 5 phonetically aware judges (Kendall's $W = .557$, $p < .001$). The degree of perceived atypicality according to group showed highly significant differences for all, with

HFA>AspS>TD, see $G-H$ values in Table 1. Difference was least between the TD and AspS groups. The mean atypicality judgments of the 5 judges correlated highly with the mean of output task scores (Pearson's $r = .57$, $p < .001$).

Among the 22 naïve judges, inter-judge agreement on atypicality was again highly significant ($W = .288$, $p < .001$), as was the perceived difference between HFA and TD groups. No difference was perceived between HFA and AspS groups, and between AspS and TD groups there was a trend towards a perception of more atypicality in the AspS group which did not quite reach significance. Correlation between the atypicality ratings and PEPS-C scores was significant ($p < .015$, $r = .612$).

Judges	Inter-judge reliability	Correlations with PEPS-C	Differences between groups		
			HFA : TD	HFA : AspS	AspS : TD
phonetically aware	$W = .557$	$r = .57$	$G-H = -0.189$	$G-H = -0.129$	$G-H = -0.059$
	$N = 5$	$N = 5$	$N = 54$	$N = 58$	$N = 54$
	$p = .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$
phonetically naïve	$W = .288$	$r = .612$	$HSD = -0.124$	$HSD = -0.034$	$HSD = -0.089$
	$N = 22$	$N = 22$	$N = 15$	$N = 15$	$N = 15$
	$p < .001$	$p = .015$	$p < .005$	ns	$p = .051$

Table 1 : Inter-judge reliability, correlation of mean atypicality with mean of PEPS-C scores, and perceived differences of atypicality between groups.

Discussion

Perception judgments correlated significantly with PEPS-C output task scores, and although the phonetically naïve judges rate the atypicality of the AspS group nearer the HFA group than the TD group, judgments broadly reflect the ranking of prosodic ability seen in Figure 1.

These results suggest that children with HFA may not only sound atypical but also be misleading or at least ambiguous in their use of prosody, although it must be borne in mind that no measure has been taken to see how far prosodic functionality in a test environment reflects prosodic functionality in a non-test situation. It also appears that, as well as sounding less atypical, the functional prosodic skills of the children with AspS are better than those of children with HFA, which raises the possibility that expressive prosody may be a distinguishing feature of the two types of ASD, perhaps as an extension of the difference between the two groups in their verbal skills.

References

- [1] Kanner, L., 1943. Autistic disturbances of affective contact. *Nervous Child*, 2, 217-250.
- [2] Fay, W.; & Schuler, A.L., 1980. *Emerging Language in Autistic Children*. London: Edward Arnold.
- [3] 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 & Phonetics*, 17(4/5), 345-354.
- [4] McCann, J.; Peppé, S., 2003. Prosody in autistic spectrum disorders: a critical review. *International Journal of Language and Communication Disorders*, 38(4), 325-350

A limitation of the research is that the naïve judges rated only a subset of the participants: further experiments will be conducted to obtain judgments from naïve judges on all the participants with ASD (and from a sample of children with TD). Additionally, since children with AspS are (as the ratings of the phonetically naïve judges suggest) often perceived to have odd prosody, the question arises as to what aspects of prosody, voice, speech or language causes this perception.

Conclusion

So far, results of the perception experiments support the ecological validity of the PEPS-C and suggest that atypicality as perceived is an indicator of prosody that may be malfunctioning, i.e. communicatively misleading. This indicates potential for the use of the PEPS-C (suitably modified) as a clinical tool for the identification of prosodic disorder, with possible implications for the classification of ASD as well as for intervention that targets prosodic skills.

- [5] Raven, J.; Court, J.; Raven, J., 1986. *Raven's Progressive Matrices and Raven's Coloured Matrices*. London: H.K.Lewis.
- [6] Dunn, L.; Dunn, L.; Whetton, C.; Burley, J., 1997. *The British Picture Vocabulary Scale (2nd edition)*. Windsor: NFER-Nelson.
- [7] Campbell, T, F.; Dollaghan, C., 1992. A method for obtaining listeners' judgements of spontaneously produced language: Social validation through direct magnitude estimation. *Topics in Language Disorder*, 12/2, 42-45.