PROSODIC BOUNDARY IN THE SPEECH OF CHILDREN WITH AUTISM

S.J.E. Peppé

Queen Margaret University, Edinburgh, UK
speppe@qmu.ac.uk

ABSTRACT

Expressive prosody is thought to be disordered in autism, and this study sets out to evaluate one aspect (prosodic boundary) to investigate a) how naïve judges rate utterances for atypicality; b) whether pitch and duration measurements in those utterances differ from those of typically-developing children; and c) whether children with autism can use prosodic boundary in speech for linguistic distinctions. Samples were drawn from children aged between 5 and 13 years; 31 with language-delayed high-functioning autism (LD-HFA), 40 with Asperger's syndrome (AS) and 119 with typical development (TD). Results showed that naïve judges perceived children with LD-HFA as sounding more atypical than those with AS, who in turn were marginally more atypical than those with TD. Measurements suggested those with LD-HFA had wider pitch-span than those with TD. The groups did not differ on linguistic functionality, and it is possible that factors other than prosody contributed to the perception of atypicality.

Keywords: atypical prosody, autism, prosodic boundary, fundamental frequency, duration.

1. INTRODUCTION

People with autism may be verbal or non-verbal, depending on the level of their autism; those who are verbal are often described as having high-functioning autism. Prosody has been noted as atypical in such individuals: this was attested first by those who originally identified autism [9, 2] and since then by many others (e.g. [14]). Atypical prosody in autism has not, however, received a great deal of attention in research, and has been labelled with adjectives that are vague (e.g. 'bizarre') and sometimes contradictory (e.g. 'monotonous' as well as 'exaggerated'). It is important to investigate whether atypical expressive prosody is misleading, thus affecting ability to communicate; or merely unusual-sounding, with a possible effect on social interaction.

Most of the previous studies of prosody in high-functioning autism have focused on stress/accent production and show inconclusive results. The focus of the current paper is to assess another aspect of prosody, the expression of boundary, in a large cohort of children with high-functioning autism and controls in three different ways: for its functionality (using a new test of functional prosodic skills); for how it is perceived by 'lay' people (i.e. adults without phonetic training); and for its acoustic properties.

We attempt to answer the following research questions:

- how children with high-functioning autism compare with controls on the ability to make functionally effective phrase-breaks;
- how consistently children with high-functioning autism are perceived as having atypical expressive prosody;
- whether there are between-group differences on measures of pitch level and pitch span;
- how far the different use of pitch and duration values affect the intelligibility of prosodic boundary;
- what pitch and duration features affect the perception of atypicality.

2. METHODS

Definitions of autism [1, 15] suggest that high-functioning autism may be divided into two kinds. Children are described as having Asperger's syndrome (AS) when they do not show clinically significant pre-school language delay. Where such delay is present, children show markedly less good language skills and may be described as having language-delayed high-functioning autism (LD-HFA). This study, which involves an aspect of language (prosody), distinguishes between the two.

2.1. Participants

Children with both conditions in the age range 6-13 years participated: 40 with AS and 31 with LD-

HFA were included. Verbally-matched typically-developing children (n = 119) took part as controls.

From each group, 25 children were randomly selected for perceptual evaluation of their prosody. The judges were 60 undergraduates without phonetic training and without experience of autistic speech.

A further 15 naïve judges provided a measure of agreement on the prosodic meaning conveyed in the data used for acoustic analysis.

2.2. Procedures

All children completed a standardised receptive vocabulary test [6] as a measure of verbal ability; and a new test of prosodic skills [12], assessing receptive and expressive ability in parallel tasks that involve the meanings conveyed by prosody: the relevant subtest for this study is the production of prosodic boundary (Chunking Expression subtest). Additionally, children in the autism groups completed a comprehensive battery of language assessments including an articulation test [8].

The experiments for perceiving atypicality of prosody involved two types of sample. One was the elicited utterance 'chocolate-biscuits and jam', a response from the PEPS-C Chunking Expression subtest: the uniformity of samples from each participant ensured that judgments would not differ according to lexical content. Another sample type was free speech, defined as 10 seconds' duration of utterances generated by a child in the course of testing. To ascertain whether perceived atypicality depends on prosody rather than on voice/lexical content, the conversation sample was low-pass filtered using PRAAT [3] with a ceiling of 500 Hz: in this sample type the prosody was audible but the lexical content unintelligible. The 60 naïve judges were asked to rate each utterance in terms of 'oddness' (atypicality), fluency, singsongness and rate; preliminary investigations had indicated that there was reasonable agreement as to what these terms meant. Judgments were made by reference to a modulus utterance using direct magnitude estimation [4].

To measure the acoustic differences, we used a subset of four other responses from the PEPS-C task ('boundary items'). In this task, children were asked to say what they saw in picture-strips depicting either three items (for example, 'cream', 'buns' and 'chocolate') or two (e.g. 'cream-buns' and 'chocolate'), each set of words suggesting different picture-strips depending on the presence or absence of phrase boundary after the first word. The possibility of ambiguity was not made explicit to the children, in order to make the task as naturalistic as possible; all, however, had previously completed a receptive task involving the same distinctions and some may have realised the need for 'meta-prosodic' control to make the distinction clear. Of responses in the Chunking task, 30 such minimal pairs from each group were suitable, and allowed us not only to measure fundamental frequency (f0) and duration in the utterance but also to rate their communicative effectiveness (how clearly the phrasing distinction was made). For the former, we took the following measurements, using PRAAT:

- Hz value at 3 points in each word in both utterances (excluding 'and'):
  - at onset of vowel (vowel-start), after microperturbation of previous segment;
  - halfway through vowel (mid-vowel);
  - at offset of vowel (vowel-end), before microperturbation of following segment;
- time (in ms):
  - at the start of word (start)
  - at the middle of the word (midpoint)
  - and at the end of word (end)

These give indications of:

- pitch-level (the final f0 point of a speaker's utterances was taken as an indicator of this, following [11]).
- pitch-span (difference in Hz between lowest and highest f0 of a speaker's total utterances).
- pitch reset between phrases, also deemed to accompany prosodic phrase-break [5]: difference (in Hz) calculated by subtracting f0 of vowel-end of the first word from f0 of vowel-start of the second word
- frequency of simple and complex on-syllable pitch-change, as in the British nuclear tone school [5] calculated by subtracting f0 of vowel-start from f0 of vowel-end (simple tones) and f0 of mid-vowel from f0 of vowel-end (complex tones).
- duration in ms of each word (including 'and') and of any pauses: end time subtracted from start time.

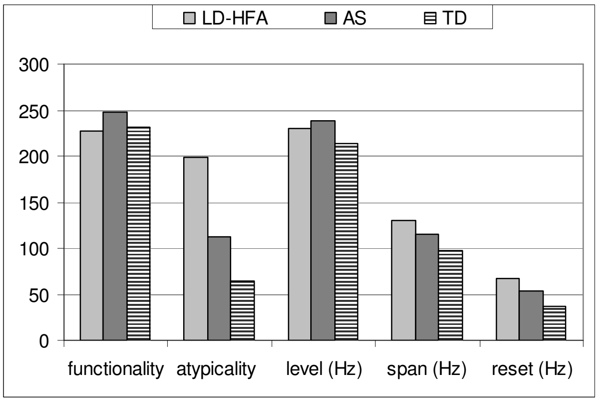
To judge communicative effectiveness, we conducted a perception experiment on the clarity of meaning of each of the four boundary items. Using a 6-point Likert scale, the 15 naïve listeners judged whether there was a prosodic boundary between the first and second word of each utterance. This made it possible to calculate a mean rating for the clarity for each utterance.

Analyses used included t-tests, Pearson's correlations and ANOVAs with Bonferroni adjustment. For the acoustic measurements, analyses were done by means of a MANOVA with group (three levels: LD-HFA, AS, and TD), age and target (two levels: 1. presence and 2. absence of boundary between the first two words).

3. RESULTS

The LD-HFA group, with a mean chronological age of 9.7 as opposed to 8.8 for the AS group and 7.7 for the TD group, showed (as expected) a significantly lower verbal mental age than the AS group (t = 2.87, p =.005) and the TD group (t = 3.49, p =.001). ANOVAs showed however no significant between-group difference on the Chunking Expression subtest scores, nor on mean judgments of functional meaning in the boundary items subset (see Figure 1).

Figure 1. Functionality, atypicality, level, span, and reset.



For atypicality in the elicited utterance, however, there was a main effect of group (F =6.38, p =.002) and also between the LD-HFA group and the other two groups (LD-HFA:AS mean difference = 86, p=.010; LD-HFA:TD mean difference = 131, p=.001; see Figure 1). The AS:TD mean difference was not significant. Perception of atypicality in conversation samples and the elicited utterance correlated highly significantly (r = .45, p <.001), but there was no significant correlation between filtered conversation samples and either the elicited utterance or the conversation samples.

In the data as a whole there was significant correlation between the perception of atypicality in the elicited utterance and the Chunking subtest scores: r =.259, p=.005; but in individual groups this correlation did not reach significance. There was also no significant correlation between perceived atypicality and mean judgments on the functional meaning of the boundary items subset, either taken as a whole or as individual groups.

ANOVAs on acoustic measurements showed some significant between-group differences, all concerning pitch measures, i.e. level, span and reset, rather than duration (see Figure 1).

The difference in pitch-level was significant between the AS and TD groups (AS higher, mean difference 24.58 Hz, p =.035), and the difference in pitch span was significant between the LD-HFA and TD groups (LD-HFA wider, mean difference 33.51 Hz, p =.013). Of the composite measures, there was a significant group difference on reset (LD-HFA greater than TD, mean difference of 29.25 Hz, p= .012). There was a significantly greater incidence of fall-rise on the first word in the TD group: LD-HFA:TD p=.037; AS:TD p=.025; and more rises on the second word in the autism groups than in the TD group (LD-HFA:TD jns, p=.051; AS:TD p=.025).

4. DISCUSSION

The analyses suggest that the functionality of this aspect of prosody in children with both types of autism is not impaired, i.e. their linguistic communication with regard to prosodic boundary is not deficient compared with their typically-developing peers. With regard to atypicality, however, ratings suggest that, as might have been expected, the LD-HFA group utterances sound most atypical, AS group utterances less so, and least atypicality is observed in the TD group.

The differences in the acoustic measurements suggests that, in this sample at least, the children with autism, especially those in the LD-HFA group, might be described as having 'exaggerated' prosody, since the pitch-span of their utterances is wider than those of the TD group. Since, however, greater pitch-span often denotes greater involvement, it may have been that the children with autism were concentrating harder, and this may have contributed to their achieving similar functional scores. The lower incidence of fall-rises is also an interesting feature, since this pattern, common in English although less so in other languages, is slow to develop in typically-developing children [13].

Apart from these findings, however, it must be borne in mind that there was no correlation between filtered samples and unfiltered ones, suggesting that prosody may not have been the main basis for judgments of atypicality. It is worth considering that, as listeners were (intentionally) not asked to specify the nature of the atypicality they observed, they may have been focusing on non-prosodic factors. One such non-prosodic factor is atypical articulation: in children with autism there is a widely held notion that articulatory ability is preserved, and indeed as a whole the groups were relatively unimpaired on the articulation test. Nevertheless, closer evaluation than that provided by the articulation test suggested that 20% of the LD-HFA group showed atypical substitutions and deviant phonological processes [7]. There was also strong correlation between articulation scores and perceived atypicality (r = .765, p <.001) in the LD-HFA group (not significant in the AS group). Another factor may be vocal quality: a preliminary voice study of children with autism [10] suggests that voice may be disordered in autism, and it was noted that one of the speakers rated highest for atypicality demonstrated a harsh vocal quality.

5. CONCLUSION

In summary, the notion that prosody is disordered in autism is borne out by the perceptual data, but the disorder seems to have little effect on the ability to make prosodic boundaries appropriately. The perception appears to have some basis in prosodic-acoustic evidence, but may depend more on factors other than prosody.

6. REFERENCES

[1] American Psychiatric Society 1994. *Diagnostic and Statistical Manual for Mental Disorders* 4th ed.. Washington, D.C: American Psychiatric Publishing, Inc.

[2] Asperger, H. 1944. Die 'Autistischen Psychopathen' im Kindesalter, *Archiv fur Psychiatrie und Nervenkrankheiten,* **117**, 76-136.

[3] Boersma, P., Weenink, D. 2004. PRAAT: doing phonetics by computer, v.4.2.06. http://www.fon.hum.uva.nl/praat/

[4] 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.

[5] Cruttenden, A. 1997 *Intonation. 2nd edition.* Cambridge: CUP

[6] Dunn, L., Dunn, L., Whetton, C., Burley, J. 1997. *The British Picture Vocabulary Scale, 2nd edition.* Windsor: NFER-Nelson.

[7] Gibbon, F., McCann, J., Peppe, S., O'Hare, A., Rutherford, M. 2004 Articulation disorders in children with high functioning autism. *Proc. 26th International Association of Logopedics & Phoniatrics Congress*, Brisbane, Australia.

[8] Goldman, R., Fristoe, M. 2000. *Goldman Fristoe-2 Test of Articulation.* Circle Pines, MN: American Guidance Service.

[9] Kanner, L. 1943. Autistic disturbances of affective contact. Nervous Child, 2, 217-250.

[10] McKinnon, A. 2005. Vocal profiling in children with language-delayed high-functioning autism. Unpublished MSc dissertation, Queen Margaret University, Edinburgh.

[11] Patterson, D. 2000. *A linguistic approach to pitch range modelling.* PhD dissertation, University of Edinburgh.

[12] 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.

[13] Wells, B., Peppé, S., Goulandris, A. 2004. Intonation development from five to thirteen. *J. Child Language* 31, 749-778.

[14] Wing, L. 1991. Asperger's syndrome and Kanner's autism. In: Frith, U. (ed), *Autism and Asperger Syndrome*. Cambridge: Cambridge University Press.

[15] World Health Organisation 1993. *The ICD-10 Classification of Mental and Behavioural Disorders.* Geneva: World Health Organisation.

.