Research in Autism Spectrum Disorders 7 (2013) 845–857

Contents lists available at SciVerse ScienceDirect

Research in Autism Spectrum Disorders

Journal homepage: http://ees.elsevier.com/RASD/default.asp

Attention to emotional tone of voice in speech perception in children with autism

Patricia J. Brooks *, Bertram O. Ploog *

College of Staten Island and the Graduate Center of City University of New York, United States

ARTICLE INFO

Article history:
Received 17 January 2013
Received in revised form 12 March 2013
Accepted 13 March 2013

Keywords:
Affective prosody
Selective attention
Speech perception
Video game
Discrimination-choice procedure

ABSTRACT

A video game was developed to assess speech perception in 13 children with autism spectrum disorders (ASD) and 13 children with typical development (TYP), ages 5–17 years old. Children listened to pre-recorded sentences varying in content (e.g., "Bob parked a van" vs. "Tim shut a door") and prosody (i.e., enthusiastic vs. grouchy tone of voice). During training, children learned to select one of two sentences differing in both content and prosody (e.g., enthusiastic "Bob parked a van" vs. grouchy "Tim shut a door"). At testing, children listened to test probes comprising re-combinations of the content and prosodic features of the training sentences. Testing indicated that both groups showed accurate discrimination of the training sentences from the re-combined test probes. However, whereas TYP children showed a preference to select the sentence with enthusiastic prosody over its grouchy counterpart, children with ASD did not. Thus, children with ASD show atypical attention to emotional tone of voice, even though they show no deficit in perceiving prosody.

To discern a speaker's communicative intentions, a listener must pay attention not only to what the speaker has said (i.e., to the content of the utterance) but also to how it has been said (i.e., to the prosody of the utterance). The ability to make sense of any indirect speech act or non-literal expression—such as a joke, hint, sarcastic statement, or metaphor—crucially hinges on an ability to use tone of voice and other prosodic and contextual cues to reject the literal meaning of an utterance and to consider an alternative meaning as pragmatically appropriate. Numerous studies report that children with autism spectrum disorders (ASD) have difficulties using speech prosody to identify emotions (Lindner & Rosén, 2006; Rutherford, Baron-Cohen, & Wheelwright, 2002; Van Lancker, Cornelius, & Kreiman, 1989) and in processing non-literal expressions (cf. Loukusa & Moilanen, 2009, for a review of studies involving individuals with Asperger Syndrome or high-functioning autism). Difficulties in identifying the emotional significance of speech prosody appear to be linked to deficits in theory of mind (Rutherford et al., 2002) as well as receptive and expressive language abilities (Brennand, Schepman, & Rodway, 2011; McCann, Peppé, Gibbon, O'Hare, & Rutherford, 2007). However, an unanswered yet fundamental question is whether such difficulties might be due in part to abnormal patterns of attention to speech prosody as opposed to some perceptual deficit, for example.

The prior literature on the perception of speech prosody in children with ASD has yielded equivocal findings, with some reporting superior performance in children with ASD relative to controls (Järvinen-Pasley, Wallace, Ramus, Happé, & Heaton, 2008), others reporting worse performance (Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2007), and others reporting

* Corresponding authors at: Department of Psychology, 4S-108, College of Staten Island, City University of New York, 2800 Victory Blvd., Staten Island, NY 10314, United States. Tel.: +1 718 982 3793.
E-mail addresses: patricia.brooks@csi.cuny.edu (P.J. Brooks), bertram.ploog@csi.cuny.edu (B.O. Ploog).

1750-9467/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.rasd.2013.03.003

(nearly) equivalent levels of performance (Chevallier, Noveck, Happé, & Wilson, 2009; Grossman, Benis, Skwerer, & Tager-Flusberg, 2010; Paul, Augustyn, Klin, & Volkmar, 2005). Recent studies utilizing neuro-imaging techniques suggest that individuals with ASD utilize a different cortical network for processing speech prosody than controls (participants with typical development, TYP), with greater involvement of cortical areas associated with attention allocation, cognitive control, sensory gating, and categorization (e.g., Eigsti, Schuh, Menci, Schultz, & Paul, 2012; Hall, Szechtman, & Nahmias, 2003). Unfortunately, a major limitation of the existing literature has been its almost exclusive focus on high-functioning children (often diagnosed with Asperger Syndrome) who are capable of grasping complex task requirements and verbal instructions. Very little is known about the processing of the prosodic features of spoken sentences in low-functioning children with ASD who may lack functional language skills.

To address the lack of inclusion of low-functioning children with ASD in extant research, Ploog, Banerjee, and Brooks (2009) embedded a discrimination-choice procedure in a video game that non-verbal children with ASD could play with success. The video-game paradigm did not require any spoken instructions on how to play the game. Instead, the children were shown manually how to manipulate the mouse to advance the game. That is, patterns of correct responding (i.e., how to receive rewards) were learned through trial and error.

In their initial investigation of how children with ASD allocate attention to speech prosody, Ploog and colleagues presented spoken sentences differing in content and intonation (i.e., with the rising intonation of a question vs. the falling intonation of a statement). During training, children were taught through trial and error to choose one of two spoken sentences differing in both content and prosody, for example, the statement "Max ate a grape!" (S+) paired with the question "Bob parked a van?" (S−). The children were rewarded for choosing the S+ according to a variable interval (VI) schedule of reinforcement and successfully learned the discrimination. Once a child reached a learning criterion of at least 75% S+ choices, test trials were used to assess the child's attention to both the content and the intonation of the sentences. Thus, for the example set of stimuli described above, recombining the features of the S+ and S− resulted in the test probes "Max ate a grape?" and "Bob parked a van!" which were individually paired with the S+, and with each other, in the test trials. Results indicated that children with ASD attended to both the content and the intonation of the sentences as the children were above chance in selecting the S+ "Max ate a grape!" when it was paired with a test probe differing only in content (e.g., "Bob parked a van!") or a test probe differing only in intonation (e.g., "Max ate a grape?") The results were similar for the TYP children with one critical exception: When the two test probes were pitted against each other (e.g., "Max ate a grape?" vs. "Bob parked a van!"), the TYP children overwhelmingly selected the test probe with the content of the S+ but differing in intonation (i.e., "Max ate a grape?"), whereas the children with ASD did not show such a preference. Thus, the results suggested no deficit in the *perception* of speech intonation in children with ASD, but rather showed a deficit of *prioritization* of attention to the content words of sentences.

In the present study, we adopted the video-game methodology of Ploog et al. (2009) to examine patterns of attention to emotional tone of voice. Perception of vocal emotion is known to be an area of weakness for children with ASD (Rutherford et al., 2002), but these children's difficulties may hinge on an inability to identify the meaning (emotional content) associated with a specific tone of voice, rather than an inability to attend to tone of voice per se (Boucher, Lewis, & Collis, 2000). Thus, the goals of our study were (1) to replicate Ploog et al. (2009) by showing that the children with ASD have no critical perceptual deficit per se but, relative to TYP children, show atypical patterns of attention; (2) to expand Ploog et al. (2009) by utilizing stimuli that differ in affective prosody (enthusiastic vs. grouchy) rather than grammatical prosody (questions vs. statements); and (3) to expand the testing procedure in that test probes were also contrasted with S−, not only S+ as in the original study. These additional test trials potentially provide information about inhibitory processes, that is, learning that has occurred with respect to stimuli predictive of the absence of reward.

1. Method

1.1. Participants

Ten boys and three girls with ASD (ages 5, 0 to 17, 6; $M = 9, 0$; $SD = 3, 4$) and six boys and seven girls with TYP development (ages 5, 4 to 16, 0; $M = 9, 0$; $SD = 3, 5$) participated. In addition, two children with ASD were recruited but did not complete participation due to an apparent lack of interest. The diagnosis of ASD was given based on school or clinic records. These children had various levels of language skills, and included non-verbal children. They were recruited from a local school for children with ASD, through a clinic at the New York State Institute for Basic Research in Developmental Disabilities, and through contacts with speech pathologists. The TYP children were recruited from an afterschool program at a campus-based school and through word of mouth.

Group assignment was strictly based on diagnosis. Matching was performed according to chronological age, not mental age. Matching by mental age has been criticized specifically for investigations of selective attention such as the current study (e.g., Gersten, 1980, 1983) because the construct of mental age remains poorly defined. Also, Ploog (2010b) pointed out that correlations between mental age and selective attention are not reliable across several studies. Nevertheless, the current data are amenable to estimating a putative effect of mental age by looking for response patterns within each group, assuming that chronological age and mental age correlate within each group (i.e., chronologically younger children also are of younger mental age). There was, however, no evidence of such an effect.