Brain & Development 36 (2014) 516–522

BRAIN & DEVELOPMENT
Official Journal of the Japanese Society of Child Neurology

www.elsevier.com/locate/braindev

Original article

Speech intonation in children with autism spectrum disorder

Yasushi Nakai [a,b,*], Ryoichi Takashima [c], Tetsuya Takiguchi [c], Satoshi Takada [a]

[a] Graduate School of Health Sciences, Kobe University, Kobe, Japan
[b] Department of Nursing Childcare, Kawasaki College of Allied Health Professions, Kurashiki, Japan
[c] Graduate School of System Informatics, Kobe University, Kobe, Japan

Received 5 February 2013; received in revised form 12 July 2013; accepted 12 July 2013

Abstract

Objective: The prosody of children with autism spectrum disorder (ASD) has several abnormal features. We assessed the speech tone of children with ASD and of children with typical development (TD) by using a new quantitative acoustic analysis. *Methods*: Our study participants consisted of 63 children (26 with ASD and 37 with TD). The participants were divided into 4 groups based on their developmental features and age. We assessed the variety of the fundamental frequency (F0) pattern quantitatively, using pitch coefficient of variation (CV), considering the different F0 mean for each word. *Results:* (1) No significant difference was observed between the ASD and TD group at pre-school age. However, the TD group exhibited significantly greater pitch CV than the ASD group at school age. (2) In pitch CV, range and standard deviation of the whole speech of each participant, no significant differences were observed between the type of participants and age. (3) No significant correlation was found between the pitch CV of each word and the Japanese Autism Screening Questionnaire total score, or between the pitch CV of each word and the intelligence quotient levels in the ASD group. A significant correlation was observed between the pitch CV of each word and social reciprocal interaction score. Conclusions: We assessed the speech tone of children with ASD by using a new quantitative method. Monotonous speech in school-aged children with ASD was detected. The extent of monotonous speech was related to the extent of social reciprocal interaction in children with ASD.

Keywords: Autism spectrum disorder; Intonation; Acoustic analysis; Pitch; Coefficient of variation; ASQ

1. Introduction

Speech prosody refers to the qualitative aspects of speech, such as intonation, rhythm, and stress [1]. The prosody of children with autism spectrum disorder (ASD) has several abnormal features. In particular, speech intonation in ASD is monotonous, which has been used as a diagnostic feature of ASD [2,3]. Speech intonation is expressed as a pitch pattern [1], using acoustic analysis, a pitch pattern can be described as a fundamental frequency (F0) pattern. The F0 is defined as the lowest frequency of a periodic waveform, and it is the physical quantity of a pitch. Therefore the variety of intonation can be interpreted as the variety of the F0 pattern.

In children with typical development (TD), the expressive abilities of falling tones have a different process pattern from those of rising tones. The range of falling tones of 4-year-old children is twice as wide as that of 1-year-old children with 4-year-old children being able to express words with falling tones as adults do [4,5]. It is more difficult for young children to express words with rising tones [6]. Crystal [7] reported that children start to learn rising tones after acquiring the ability

* Corresponding author. Address: Department of Nursing Childcare, Kawasaki College of Allied Health Professions, 316 Matsushima, Kurashiki 701-0194, Japan. Tel.: +81 86 464 1032; fax: +81 86 463 4339.
E-mail address: nakai@jc.kawasaki-m.ac.jp (Y. Nakai).

0387-7604/$ - see front matter © 2013 The Japanese Society of Child Neurology. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.braindev.2013.07.006

to use falling tones. There is no significant difference in the variety of rising tones between 1-year-old children and 4-year-old children [4], which supports the idea that children begin to learn rising tones after they have reached school age [8].

Monotonous speech in ASD was first described by Kanner, and has since been consistently observed in a large number of studies [9–11]; moreover, it is evident not only in childhood but also throughout adolescence and adulthood [10,12,13]. Monotonous speech is considered a core feature of ASD, yet may not be related to the extent of ASD [14]. However, all of the reports relating to monotonous speech in ASD to date were based on subjective observations. The reliability and validity of human auditory abilities are often lower than predicted [15]. Therefore, a method of quantitative assessments has been sought.

In recent years, a few studies have shown that monotonous speech in ASD is strongly linked to the impairment of speech prosody perception [13,16]. For example, abnormal processing of auditory stimuli in ASD has been speculated to occur at the brainstem level [17], and abnormalities have been revealed by functional magnetic resonance imaging (fMRI) [18]. Furthermore, speech prosody perception in ASD is considered to be relevant to the development of the Theory of Mind [19]. However, monotonous speech in ASD has not been confirmed in any former quantitative studies that used either pitch range (the minimum F0 subtracted from maximum F0) or pitch standard deviation (SD) [20–22].

The objective of this study, therefore, was to clarify the differences in speech intonation between children with ASD and children with TD. We assessed the speech tone of children with ASD by using a new quantitative method. Additionally, we sought to measure F0 using acoustic analysis, and to examine how the variety of the F0 pattern of children with ASD and of children with TD changes between pre-school and school ages. Furthermore, we assessed the relationship between the variety of the F0 pattern and the extent of ASD symptoms.

2. Methods

2.1. Participants

ASD participants were recruited from among the 4- to 9-year-old children who visited the Developmental Behavioral Pediatric Clinic of Kobe University Hospital between April and July 2010, after approval from the Medical Ethics Committee of Kobe University Graduate School of Medicine was received. All of the participants had a clinical diagnosis of ASD based on DSM-IV-TR [2] criteria made by an expert child neurologist. Participants were examined routinely by the one-foot standing test, the finger-to-nose test, and the simple drawing test during their first visit. None of the participants revealed obvious neurological symptoms such as involuntary movements. Participants and their parents were provided with a complete written and oral explanation of the aims and protocols of the study prior to providing their informed consent. They were assured that all personal information would be protected, that non-participation in the study would disadvantage them in any way, and that participants or their parents could discontinue the participation in the study at any time. Participation in the study was limited to children who were able to understand simple instructions and who had obtained 30 or more words for analysis to ensure reliability.

The control or TD participants were recruited from among the children studying in the normal class of their preschools or primary schools, which were in the same region as the ASD participants. None had any history of speech, communication, learning problems, or any history of special education.

Our study participants consisted of a total of 63 children (26 with ASD and 37 with TD). We divided the participants into 4 groups based on their developmental features and age. The 4- to 6-year-old (pre-school-aged) ASD group consisted of 6 children, and the 7- to 9-year-old (school-aged) ASD group consisted of 20 children. The pre-school-aged TD group consisted of 16 children, and the school-aged TD group consisted of 21 children (Table 1).

2.2. Procedures

We conducted experimental tasks based on the word test proposed by the Japan Clinical Articulation Society [23]. In this test, 50 picture cards of animals and objects were displayed, and participants were instructed to verbally name each object or animal. After displaying all 50 picture cards, we took a 3-min break before repeating the exercise. When picture cards were displayed, the examiners asked the participants to name the object shown. Because the examiner did not say the name of the object, the participants were not able to directly imitate their speech. If the participants gave an incorrect response or did not answer, the examiners displayed the next picture card without providing them with the correct term. The participants were told the correct terms after all word tests had been completed. Incorrect answers were excluded from final data analysis.

The recording equipment consisted of a Lavalier Microphone (Sony, ECM-77B/9X), Audio Capture (ROLAND, EDIROL UA-25EX), and a Mobile Note PC (TOSHIBA, Dynabook SS RX2/T7G). The microphone was fixed to the participant's collar during recording. Recording parameters included a 48 kHz sampling frequency and 24-bit quantization.