Research in Developmental Disabilities xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Research in Developmental Disabilities

journal homepage: www.elsevier.com/locate/redevdis

Typical vs. atypical: Combining auditory Gestalt perception and acoustic analysis of early vocalisations in Rett syndrome

Florian B. Pokorny[a,b,c], Katrin D. Bartl-Pokorny[a], Christa Einspieler[a], Dajie Zhang[a,d], Ralf Vollmann[c,e], Sven Bölte[f,g], Markus Gugatschka[a], Björn W. Schuller[h,i], Peter B. Marschik[a,c,d,f,*]

[a] iDN – interdisciplinary Developmental Neuroscience, Department of Phoniatrics, Medical University of Graz, Graz, Austria
[b] Machine Intelligence & Signal Processing group, MMK, Technical University of Munich, Munich, Germany
[c] Brain, Ears & Eyes – Pattern Recognition Initiative, BioTechMed-Graz, Graz, Austria
[d] Department of Child and Adolescent Psychiatry and Psychotherapy, iDN – interdisciplinary Developmental Neuroscience, University Medical Center Göttingen, Göttingen, Germany
[e] Department of Linguistics, University of Graz, Graz, Austria
[f] Center of Neurodevelopmental Disorders (KIND), Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden
[g] Child and Adolescent Psychiatry, Center for Psychiatry Research, Stockholm County Council, Sweden
[h] Chair of Embedded Intelligence for Health Care and Wellbeing, Augsburg University, Augsburg, Germany
[i] Group on Language, Audio & Music (GLAM), Department of Computing, Imperial College London, London, UK

ARTICLE INFO

Number of reviews completed is 2.
The paper is part of a special issue on Gene x Environment Int.

Keywords:
Rett syndrome
Preserved speech variant
Speech-language pathology
Early vocalisations
Auditory perception
Acoustic vocalisation analysis

ABSTRACT

Background: Early speech-language development of individuals with Rett syndrome (RTT) has been repeatedly characterised by a co-occurrence of apparently typical and atypical vocalisations.

Aims: To describe specific features of this intermittent character of typical versus atypical early RTT-associated vocalisations by combining auditory Gestalt perception and acoustic vocalisation analysis.

Methods and procedures: We extracted N = 363 (pre-)linguistic vocalisations from home video recordings of an infant later diagnosed with RTT. In a listening experiment, all vocalisations were assessed for (a)typicality by five experts on early human development. Listeners' auditory concepts of (a)typicality were investigated in context of a comprehensive set of acoustic time-, spectral- and/or energy-related higher-order features extracted from the vocalisations.

Outcomes and results: More than half of the vocalisations were rated as 'atypical' by at least one listener. Atypicality was mainly related to the auditory attribute 'timbre', and to prosodic, spectral, and voice quality features in the acoustic domain.

Conclusions and implications: Knowledge gained in our study shall contribute to the generation of an objective model of early vocalisation atypicality. Such a model might be used for increasing caregivers' and healthcare professionals' sensitivity to identify atypical vocalisation patterns, or even for a probabilistic approach to automatically detect RTT based on early vocalisations.

* Corresponding author at: Research Unit iDN – interdisciplinary Developmental Neuroscience, Department of Phoniatrics, Medical University of Graz, Auenbruggerplatz 26, Graz, 8036, Austria.
E-mail addresses: peter.marschik@medunigraz.at, peter.marschik@ki.se (P.B. Marschik).

https://doi.org/10.1016/j.ridd.2018.02.019
Received 30 October 2017; Received in revised form 20 February 2018; Accepted 21 February 2018

What this paper adds?

Our study adds to the state of knowledge that the early speech-language development of individuals with Rett syndrome (RTT) already bears atypicalities that are co-occurring with apparently typical verbal behaviour. Here, we provide a comprehensive delineation of this intermittent character of typical versus atypical early verbal behaviour in RTT on the basis of a comprehensive set of early RTT-associated vocalisations. To the best of our knowledge, this is the first study to describe early vocalisation atypicalities in RTT by combining auditory Gestalt perception in the framework of a listening experiment, with a precise acoustic vocalisation analysis. This allowed us to objectify listeners' auditory concepts of early vocalisation (a)typicality on the signal level. Our approach may pave the way for the generation of an objective model of early vocalisation atypicality in RTT or other developmental disorders usually diagnosed in or beyond toddlerhood. Such a model could be used for increasing the sensitivity of caregivers and healthcare professionals to identify early vocalisation atypicalities, and further to implement a probabilistic tool for an automated vocalisation-based earlier detection of RTT or other 'late diagnosed' developmental disorders.

1. Introduction

Rett syndrome (RTT, MIM 312750) is a severe progressive developmental disorder that almost exclusively affects females at a prevalence of approximately 1 in 10,000 live female births (Laurvick et al., 2006). More than 30 years after the first description of the clinical presentation of RTT by the Austrian neuropediatrician Andreas Rett (1966, 2016), mutations in the X-linked gene *MECP2* were identified as the main cause of the disease (Amir et al., 1999). There are, however, individuals with *MECP2* mutations showing no clinical signs and individuals with a RTT phenotype without *MECP2* mutations (Neul et al., 2010; Suter, Treadwell-Deering, Zoghbi, Glaze, & Neul, 2014). The latter can be, at least for a proportion, explained as variants or atypical forms of RTT related to mutations in other genes (e.g., *FOXG1*, *CDKL5*; Neul et al., 2010; Sajan et al., 2017). Thus, RTT remains a clinical diagnosis based upon four main clinical consensus criteria (Neul et al., 2010): (1) partial or complete loss of purposeful hand skills; (2) partial or complete loss of acquired spoken language; (3) gait abnormalities (dyspraxic or absent gait); and (4) stereotypic hand movements (e.g., repetitive washing-like movements, clapping movements, hand-to-mouth stereotypies). In addition to the main criteria, Neul et al. (2010) defined eleven supportive diagnostic criteria: breathing disturbances and bruxism when awake, sleep problems, abnormal muscle tone, vasomotor changes, scoliosis/kyphosis, growth retardation, small cold hands and feet, inappropriate laughing/screaming spells, reduced response to pain, and intense eye communication. An individual is diagnosed with typical RTT, if all four main criteria are met, or with atypical RTT, if at least two main criteria and five supportive criteria are met. According to these consensus criteria, we differentiate between three atypical forms of RTT, with the relatively milder preserved speech variant (PSV; Zappella variant) being one of them (Neul et al., 2010). PSV is typically associated with comparably better recovery of speech-language capacities and functional hand use after regression (Marschik, Einspieler, Oberle, Laccone, & Prechtl, 2009; Renieri et al., 2009). For parents, but also for healthcare professionals seeing the child during clinical routine examinations, early development in RTT might appear inconspicuous. Thus, the mean age of diagnosis is still at 2.7 years (Tarquinio et al., 2015). Often, the regression of already acquired functions raises parental concerns, and is the first warning signal that motivates parents to consult a specialist. There is, however, growing evidence that atypicalities in different developmental areas are already present prior to the onset of regression. Studies employing retrospective analysis of family videos of individuals who are later diagnosed with typical RTT or PSV have revealed deviances from typical development in motor functions and speech-language/socio-communicative capabilities during the first year of life (Bartl-Pokorny et al., 2013; Burford, Kerr, & Macleod, 2003; Einspieler, Kerr, & Prechtl, 2005a; Einspieler, Kerr, & Prechtl, 2005b; Einspieler, Freilinger, & Marschik, 2016; Einspieler, Marschik et al., 2014; Einspieler, Sigafoos et al., 2014; Leonard & Bower, 1998; Marschik et al., 2009, 2013; Marschik, Bartl-Pokorny et al., 2014; Marschik, Einspieler, & Sigafoos, 2012; Marschik, Kaufmann et al., 2012, Marschik, Pini et al., 2012; Marschik, Sigafoos et al., 2012; Marschik, Vollmann et al., 2014; Pokorny, Marschik, Einspieler, & Schuller, 2016; Tams-Little & Holdgrafer, 1996; Townend et al., 2015). In our previous work on aspects of early speech-language development, we have found that a proportion of individuals with typical RTT and PSV do not achieve certain early speech-language milestones (e.g., cooing, babbling, proto-words; Bartl-Pokorny et al., 2013; Marschik, Bartl-Pokorny et al., 2014; Marschik et al., 2013). Observed vocalisations occurred with an intermittent character of typical and atypical patterns (i.e., pressed, inspiratory, or high-pitched crying-like; Marschik et al., 2013; Marschik, Pini et al., 2012). In a previous listening experiment, Marschik, Einspieler et al. (2012) demonstrated that listeners were able to distinguish between the vocalisation sequences of typically developing infants and pre-selected atypical vocalisation sequences of infants later diagnosed with RTT. Interestingly, a vocalisation which was produced by an individual later diagnosed with RTT, and considered typical by speech-language pathologists, was not classified consistently in this experiment. These findings together with clinical and parental observations, indicate a need for more detailed characterisation of early verbal behaviour in RTT. To the best of our knowledge, the study presented here is the first that aims to comprehensively delineate specific features of the intermittent character of typical versus atypical early vocalisations in RTT. Our approach is unique in that it combines a listening experiment on an extensive sample of early RTT-associated vocalisations, with precise signal analytical descriptions. In this way, auditory analysis may be extended to include an acoustic perspective, shedding further light on aspects of early vocalisation development in RTT. With this study, we intended to pave the way for an approach in which the communicative environment of an infant with a not yet diagnosed genetic disorder can be actively involved in the promotion of an early identification of atypical behavioural patterns to facilitate earlier intervention. Such an approach seems to be promising as infants actively influence the communicative environment through their early verbal behaviours (cf. the concept of the self-generated environment as, for example, described by Esposito, Hiroi, and Scattoni (2017) for infants with autism spectrum disorder). In this regard, a higher amount and similar pattern of atypical vocalisations uttered by an infant might more likely lead to