

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Research in Autism Spectrum Disorders

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

Temporal synchrony and audiovisual integration of speech and object stimuli in autism



Elizabeth Smith*, Shouling Zhang, Loisa Bennetto

Department of Clinical and Social Sciences in Psychology, University of Rochester, Rochester, NY, USA

ARTICLE INFO

Keywords:

Speech processing
 Audiovisual integration
 Temporal perception

ABSTRACT

Background: Individuals with Autism Spectrum Disorders (ASD) have been shown to have multisensory integration deficits, which may lead to problems perceiving complex, multisensory environments. For example, understanding audiovisual speech requires integration of visual information from the lips and face with auditory information from the voice, and audiovisual speech integration deficits can lead to impaired understanding and comprehension. While there is strong evidence for an audiovisual speech integration impairment in ASD, it is unclear whether this impairment is due to low level perceptual processes that affect all types of audiovisual integration or if it is specific to speech processing.

Method: Here, we measure audiovisual integration of basic speech (i.e., consonant-vowel utterances) and object stimuli (i.e., a bouncing ball) in adolescents with ASD and well-matched controls. We calculate a temporal window of integration (TWI) using each individual's ability to identify which of two videos (one temporally aligned and one misaligned) matches auditory stimuli. The TWI measures tolerance for temporal asynchrony between the auditory and visual streams, and is an important feature of audiovisual perception.

Results: While controls showed similar tolerance of asynchrony for the simple speech and object stimuli, individuals with ASD did not. Specifically, individuals with ASD showed less tolerance of asynchrony for speech stimuli compared to object stimuli. In individuals with ASD, decreased tolerance for asynchrony in speech stimuli was associated with higher ratings of autism symptom severity.

Conclusions: These results suggest that audiovisual perception in ASD may vary for speech and object stimuli beyond what can be accounted for by stimulus complexity.

1. Introduction

Multisensory integration is a ubiquitous process by which information combined from multiple sensory streams leads to faster, more robust, and more accurate perception of the world around us. For example, when listening to a person speak, watching their mouth move leads to a faster auditory brain response ([van Wassenhove, Grant, & Poeppel, 2005](#)), louder-sounding speech ([MacLeod & Summerfield, 1987](#)), and increased understanding of both words ([Schwartz, Berthommier, & Savariaux, 2004](#)) and ideas ([Arnold & Hill, 2001](#); [Reisberg, McLean, & Goldfield, 1987](#)). Some aspects of audiovisual integration are already developed and robust within the first year of life both at neural ([Hyde, Jones, Porter, & Flom, 2010](#)) and behavioral levels ([Lewkowicz, 2010](#)). In fact, audiovisual integration in infancy predicts later language and communication abilities ([Kushnerenko et al., 2013](#)), which is evidence

* Corresponding author. Present Address: The Child Development Lab, Department of Human Development, 3304 Benjamin Building, University of Maryland, College Park, MD, 20742-1131, USA.

E-mail addresses: esmith23@umd.edu (E. Smith), shouling.zhang1@gmail.com (S. Zhang), loisa.bennetto@rochester.edu (L. Bennetto).

<http://dx.doi.org/10.1016/j.rasd.2017.04.001>

Received 1 December 2016; Received in revised form 27 March 2017; Accepted 2 April 2017

1750-9467/© 2017 Elsevier Ltd. All rights reserved.

for the importance of audiovisual integration in early language development. While audiovisual integration is present and important early in life, it continues to mature through childhood and adolescence (Dick, Solodkin, & Small, 2010; Knowland, Mercure, Karmiloff-Smith, Dick, & Thomas, 2014; Lewkowicz & Flom, 2014)

Mounting evidence supports an audiovisual integration impairment in ASD, although signs of this impairment may be subtle. For example, individuals with autism show a reduced McGurk effect (Mongillo et al., 2008; Woynaroski et al., 2013), which persists even when their looking time to faces is the same as in controls (Irwin, Tornatore, Brancazio, & Whalen, 2011). Individuals with autism also show reduced audiovisual benefit when listening to speech in background noise (Smith & Bennetto, 2007), although this may improve with age (Foxy et al., 2013). These differences are reflected at the neurophysiological level as shown in studies using both electroencephalography (Brandwein et al., 2013, 2014; Megnin et al., 2012) and fMRI (Hubbard et al., 2012). Notably, studies using the non-social “sound induced flash illusion” (Shams, Kamitani, & Shimojo, 2000) have shown intact integration of nonsocial stimuli in autism (i.e., presence of the illusion, Foss-Feig et al., 2010; van der Smagt, van Engeland, & Kemner, 2007).

Temporal synchronization of auditory and visual cues is a feature of audiovisual integration that plays a significant role in the degree of audiovisual benefit. Reducing audiovisual temporal synchrony can reduce both illusory audiovisual effects (Stevenson, Zemtsov, & Wallace, 2012) and audiovisual reaction time benefits (Diederich & Colonius, 2004). For any construct capturing audiovisual integration, there is thought to be a window of audiovisual asynchronies within which the dependent variable increases probabilistically, frequently referred to as the Temporal Window of Integration (TWI, Spence & Squire, 2003; van Wassenhove, Grant, & Poeppel, 2007). Recently, the effects of temporal synchrony on audiovisual integration in autism have been measured across multiple studies (de Boer-Schellekens, Eussen, & Vroomen, 2013; de Boer-Schellekens, Keetels, Eussen, & Vroomen, 2013; Foss-Feig et al., 2010; Grossman, Schneps, & Tager-Flusberg, 2009; Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011; Stevenson et al., 2014; Woynaroski et al., 2013). Findings have included a larger temporal window of integration in autism across multiple stimulus types (de Boer-Schellekens, Eussen et al., 2013; de Boer-Schellekens, Keetels et al., 2013; Foss-Feig et al., 2010; Kwakye et al., 2011; Woynaroski et al., 2013), differentially larger temporal windows for speech stimuli only (Stevenson et al., 2014), or no difference in temporal windows (Grossman et al., 2009). Importantly, TWI can be affected by stimulus type (e.g., speech, object), stimulus complexity (e.g. complex human speech vs consonant-vowel utterances), and choice of dependent variable (e.g., presence of illusions, temporal order judgment, or simultaneity judgment), which may explain these discrepancies (Stevenson & Wallace, 2013; van Eijk, Kohlrausch, Juola, & van de Par, 2008; Vatakis, Navarra, Soto-Faraco, & Spence, 2008; Vatakis & Spence, 2006, 2010). While it is known that social and nonsocial stimuli are associated with different TWIs and that individuals with autism tend to show less impairment in nonsocial audiovisual integration paradigms, only one study directly compares social and nonsocial temporal synchrony perception in autism (Stevenson et al., 2014). Thus, while differences in multisensory integration for social versus nonsocial stimuli are a hallmark of the audiovisual integration literature in autism, we know little about the role of temporal synchrony detection in this pattern.

Most audiovisual integration tasks by their nature require various degrees of verbal processing. Specifically, they may require the subject to report on their perception, stating 1) whether sound or video appeared first (temporal order judgment) 2) whether the audio and video were in synch (simultaneity judgment) 3) what speech sound they heard (McGurk effect) or 4) how many times they heard a beep (flash-beep illusion). While responding with single words is not likely to be particularly taxing for the individuals with autism participating in these studies (i.e. those without intellectual disability), reduction of verbal task demands is preferable, particularly if those studies are to be extended to individuals with more significant verbal deficits.

While there is evidence that minimizing verbal demands still leads to group differences in audiovisual integration tasks in autism (Collignon et al., 2013), it is unknown whether this is the case for temporal synchrony tasks. Beyond verbal processes related to task demands, many investigations of temporal synchrony detection and its role in audiovisual integration in autism use speech stimuli that are both more visually and auditorily complex than those used for nonspeech conditions. Because more complex stimuli are generally associated with larger temporal windows of integration (Vatakis & Spence, 2006, 2010), this could lead to finding larger TWIs in autism that are related to the complexity of the stimuli rather than its social nature (Williams, Minshe, & Goldstein, 2015). Here, we use speech and object-based audiovisual stimuli designed to minimize differences in complexities (i.e., in timing, pitch, etc) paired with a nonverbal temporal synchrony task to determine the role of stimulus type in audiovisual integration in autism. Specifically, we used a paradigm based on detection versus discrimination of temporal synchrony, and used speech and object stimuli shown in a pilot study to have similar TWIs. The task in the present study requires participants to use a physical response (i.e., pushing arrow buttons vs. verbal response) to select which of two videos matches auditory stimuli in both speech and object conditions. Selecting the matching video parallels the essential process of using visual cues to locate a speaker in a noisy room (i.e., the Audiovisual Cocktail Party Effect, (Zion Golumbic, Cogan, Schroeder, & Poeppel, 2013)). While similar designs have shown sensitivity to audiovisual temporal synchrony in autism and typical development, (Baart, Vroomen, Shaw, & Bortfeld, 2014; Bebko, Weiss, Demark, & Gomez, 2006; Patten, Watson, & Baranek, 2014), the present design incorporates subtly varying audiovisual asynchronies in order to measure the TWI for responses to both speech and object stimuli.

2. Methods

This study was approved by the University of Rochester Institutional Review Board, and all subjects and parents of subjects 17 years of age and younger completed informed consent.

Download English Version:

<https://daneshyari.com/en/article/4940888>

Download Persian Version:

<https://daneshyari.com/article/4940888>

[Daneshyari.com](https://daneshyari.com)