Auditory–Visual Misalignment: A Theoretical Perspective on Vocabulary Delays in Children With ASD

Courtney E. Venker, Allison Bean and Sara T. Kover

In this commentary, we describe a novel theoretical perspective on vocabulary delays in children with autism spectrum disorder (ASD)—a perspective we refer to as auditory–visual misalignment. We synthesize empirical evidence that: (a) as a result of differences in both social and nonsocial visual attention, the auditory–visual statistics available to children with ASD for early word learning are misaligned; (b) this auditory–visual misalignment disrupts word learning and contributes to the vocabulary delays shown by children with ASD; and (c) adopting a perspective of auditory–visual misalignment has important theoretical and clinical implications for understanding and supporting vocabulary development in children with ASD. Theoretically, the auditory–visual misalignment perspective advances our understanding of how attentional differences impact vocabulary development in children with ASD in several ways. By adopting the point of view of the child, we provide a framework that brings together research on social and domain-general visual attention differences in children with ASD. In addition, the auditory–visual misalignment perspective moves current thinking beyond how misalignment disrupts vocabulary development in the moment, and considers the likely consequences of misalignment over developmental time. Finally, considering auditory–visual misalignment may assist in identifying active ingredients of existing language interventions or in developing new interventions that deliver high quality, aligned input. Future research is needed to determine how manipulating auditory–visual alignment changes word learning in ASD and whether the effects of auditory–visual misalignment are unique to ASD or shared with other neurodevelopmental disorders or sources of language impairment. ***Autism Research*** *2018, 11: 1621–1628.* © 2018 International Society for Autism Research, Wiley Periodicals, Inc.

Lay Summary: This article describes a new way of thinking about vocabulary delays in children with autism spectrum disorder (ASD). We suggest that children with ASD may have difficulty learning words because their attention is not tuned in to what is most important for learning, creating a mismatch between what they see and what they hear. This perspective brings together research on different types of attentional differences in people with ASD. It may also help us to understand how language interventions work.

Keywords: attention; language; visual; auditory; learning; developmental psychology

Introduction

Many children with autism spectrum disorder (ASD) demonstrate delayed vocabulary development [Charman et al., 2003; Charman, Howlin, Berry, & Prince, 2004; Ellis Weismer, Lord, & Esler, 2010; Luyster, Kadlec, Carter, & Tager-Flusberg, 2008; Paul, Chawarska, Cicchetti, & Volkmar, 2008], yet the word-learning mechanisms that support lexical acquisition appear to be largely intact [Haebig, Saffran, & Ellis Weismer, 2017; Mayo & Eigsti, 2012; Naigles, Kelty, Jaffery, & Fein, 2011; Swensen, Kelley, Fein, & Naigles, 2007]. This seemingly incongruent finding of delayed vocabulary development and intact word-learning mechanisms has led to the proposal that vocabulary delays in children with ASD are due not to impaired learning mechanisms per se, but to difficulty with *intake of the input* resulting from disruptions in supporting cognitive systems, such as attention [Arunachalam & Luyster, 2016; Tenenbaum, Amso, Righi, & Sheinkopf, 2017].

However, inefficient intake of language input due to disruptions in supporting cognitive systems may not fully explain the word learning difficulties experienced by children with ASD. In the current commentary, we suggest that disruptions in one cognitive system (attention) play a critical role—not only limiting intake of the input, but *fundamentally changing the input* itself, with clear consequences for lexical acquisition. The rationale

From the Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, Michigan (C.E.V.); Department of Speech and Hearing Science, The Ohio State University, Columbus, Ohio (A.B.); Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington (S.T.K.)

Received June 27, 2018; accepted for publication September 29, 2018

Address for correspondence and reprints: Courtney E. Venker, Department of Communicative Sciences and Disorders, Michigan State University, 1026 Red Cedar Rd., Room 216, East Lansing, MI 48824. E-mail: cvenker@msu.edu

Published online 26 November 2018 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/aur.2038

is that differences in visual attention among children with ASD alter the co-occurrences these children perceive between auditory and visual information (i.e., words, and the objects, or events they describe), resulting in a misalignment of the auditory–visual statistics available for vocabulary development.

Our goal in this commentary is to describe how differences in visual attention among individuals with ASD produce a mismatch of statistical co-occurrences that we refer to as auditory–visual misalignment, and how such auditory–visual misalignment is likely to contribute to vocabulary delays in children with ASD. We do not comprehensively review scientific evidence of differences in visual attention among individuals with ASD, as others have done [Bottema-Beutel, 2016; Keehn, Muller, & Townsend, 2013; Sacrey, Armstrong, Bryson, & Zwaigenbaum, 2014]. Instead, we synthesize empirical evidence that: (a) as a result of differences in both social and nonsocial visual attention, the auditory–visual statistics available to children with ASD for early word learning are misaligned; (b) this auditory–visual misalignment disrupts word learning and contributes to the vocabulary delays shown by children with ASD; and (c) adopting a perspective of auditory–visual misalignment has important theoretical and clinical implications for understanding and supporting vocabulary development in children with ASD.

Although we fully acknowledge the importance of auditory attention in language development in children with ASD [Foss-Feig, Schauder, Key, Wallace, & Stone, 2017], the current commentary focuses on how visual attention impacts the alignment of auditory–visual statistics relevant for vocabulary acquisition.

Auditory–Visual Co-Occurrences in Early Word Learning: Theoretical Basis

Much of early word learning relies on children's ability to associate the words they hear with the objects or events they see. For example, children learn that "ball" describes a round toy and that "jumping" refers to the act of propelling oneself off the ground. Thus, the relevant input for early word learning is largely comprised of correlated auditory and visual statistics: objects/events, and the words that describe them. Although the world presents many things to see and hear, children cannot make use of all of these auditory–visual co-occurrences. Instead, children must focus on some things to the exclusion of others, thereby constraining the statistics to which they are exposed [Smith, Suanda, & Yu, 2014]. Children, therefore, play a key role in constructing the auditory–visual statistics they experience, because they control where they direct their visual attention at any given time [Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006; Smith & Yu, 2008; Tenenbaum et al., 2017].

When children's visual attention is tuned in to what is relevant for a given learning opportunity, they accumulate auditory–visual statistics that are aligned (i.e., related to one another). For example, a child's ability to follow an adult's point and verbal cue to, "Look at the doggie!" produces alignment between what the child sees (the doggie) and what the child hears (e.g., "Look at the doggie! He's running so fast!"). As a result of this auditory–visual alignment, the child has an opportunity to either learn new words or strengthen existing representations. The role of auditory–visual alignment in vocabulary development is important to consider because, as outlined by the developmental-dynamic view of word learning, when a word and an object/event co-occur, associative learning mechanisms strengthen the link between them [Kucker, McMurray, & Samuelson, 2015; Tamis-LeMonda, Kuchirko, & Song, 2014]. Indeed, temporal co-occurrences plays a major role in determining the probability that two pieces of information will be associated [Rovee-Collier, 1995; Tamis-LeMonda et al., 2014]. In this way, looking at the right thing at the right time allows children to build up correct associations between words and their referents, and prune away incorrect or inconsistent ones, which supports learning in the moment and gradually produces a functional vocabulary [Kucker et al., 2015].

Auditory–Visual Alignment and Language in Typical Development

Although it is not usually described using this terminology, there is strong empirical evidence that auditory–visual alignment supports vocabulary development in children with typical development (and that auditory–visual *mis*alignment disrupts it). Much of this evidence comes from observational studies demonstrating a link between young children's language, and the extent to which they initiate or respond to bids for shared, object/event-focused attention. In typical development, young children often follow adult eye-gaze, points, or verbal cues, or produce these cues themselves to direct people to the things that interest them—a set of related behaviors referred to, respectively, as responding to joint attention (RJA) and initiating joint attention [IJA; Bakeman & Adamson, 1984; Brooks & Meltzoff, 2005; Mundy & Gomes, 1998; Tomasello & Farrar, 1986]. From an alignment perspective, joint attention behaviors are important because they create many instances in which adults and children are looking at (and likely talking about) the same aspects of the world—in other words, auditory–visual alignment.

If auditory–visual alignment is beneficial, then interacting with the world in a way that produces an accumulation of aligned visual–auditory statistics should support vocabulary development. In line with this argument, children who have better joint attention skills early in life have better language skills later in life. Even before the age of two, toddlers' ability to follow adult gaze is positively associated with language abilities [Brooks & Meltzoff, 2005; Mundy & Gomes, 1998]. Observational work using head cameras also provides supporting evidence: when parents label novel objects during naturalistic play interactions, the objects children learn the labels for are the objects they were looking at before, during, and after the label was presented [Pereira, Smith, & Yu, 2014].

There is also growing evidence from experimental word-learning studies that auditory–visual alignment helps children link new words and their meanings. Axelsson, Churchley, and Horst [2012] found that word learning in 24-month olds with typical development was enhanced when children's visual attention was focused on target objects, rather than on competitors. Based on these findings, the authors suggested that sustained attention on a target object, "facilitates the processing and encoding of information about the novel object, the novel name, and their association, such that this information can be recalled after a delay" (p. 7). Similarly, Pomper and Saffran [2018] found that typically developing 3-year olds' ability to identify named referents, as well as learn and retain word meanings, was disrupted when a salient competitor object pulled children's attention away from the target object. For typically developing children, it appears that looking at the right thing at the right time facilitates word learning, whereas looking the wrong thing can be detrimental. What does this mean for children with ASD, whose differences in visual attention are likely to produce relatively high rates of auditory–visual misalignment?

Auditory–Visual Misalignment and Language in Children with ASD

Social Attention

Children with ASD are less likely than children with typical development to engage in shared, object/event-focused attention [Bottema-Beutel, 2016; Kasari, Sigman, Mundy, & Yirmiya, 1990; Mundy & Jarrold, 2010; Mundy, Sigman, & Kasari, 1990]. Thus, they are more likely to experience situations in which they hear a label for one object or event, while looking at another—auditory–visual misalignment. For example, imagine that the child in our earlier example had been unaware of, had misunderstood, or had been unwilling or unable to follow the adult's cues to look at the running dog. Instead of receiving aligned auditory–visual statistics, the child would have received auditory input that was *not* aligned with what he was seeing at that moment (e.g., a sitting cat). This auditory–visual misalignment may have prevented him from learning new words or strengthening their representations, and may even have resulted in incorrect learning—which, over time, could decrease this child's motivation for vocabulary acquisition [Baron-Cohen, Baldwin, & Crowson, 1997].

In addition to disrupting learning in the moment, auditory–visual misalignment may disrupt vocabulary development over the long term, as the child was exposed to many spurious (i.e., incorrect or irrelevant) associations between words, and objects or events. Consistent with this proposal, there is empirical evidence that auditory–visual misalignment resulting from differences in social attention can disrupt vocabulary learning and may have a cascading impact on language development in the long term. Just as in typical development, a stronger ability to create alignment (by exhibiting RJA and IJA) is associated with better language abilities in children with ASD [Bono, Daley, & Sigman, 2004; Bottema-Beutel, 2016; Kasari, Gulsrud, Freeman, Paparella, & Hellemann, 2012]. Furthermore, joint attention intervention can lead to gains in language skills in young children with ASD, particularly in children with the most severe initial language delays [Kasari, Paparella, Freeman, & Jahromi, 2008].

Perhaps some of the clearest evidence that misalignment can be problematic comes from an experimental study by Baron-Cohen et al. [1997]. In this study, children with ASD learned new words correctly when the experimenter labeled objects within the children's focus of attention. However, the majority of children made mapping errors when the experimenter labeled objects only within the experimenter's focus of attention. In such cases, children associated the novel words within their own focus of attention, rather than the object the experimenter was looking at and thus had intended to label. These findings provide evidence that misalignment can lead to incorrect learning in children with ASD. Although additional empirical evidence of incorrect word learning in children with ASD is scarce, idiosyncratic word usage (tracked through anecdotal reports) may also be the product of incorrect learning. For example, Baron-Cohen et al. [1997] described a child who referred to a toy truck as a "sausage." "In this case, the child's mother explained that the toddler had been playing with a toy truck when the mother said, 'Tommy, come and eat your sausage.' At the time, the child had been looking at his toy truck on the floor while the mother was facing away, looking at the plate of food on the table. The child had presumably failed to check the mother's direction of gaze, and so had learned the wrong association" (p. 50).

In addition to potentially producing incorrect learning, auditory–visual misalignment may result in the absence

of learning—a point illustrated by Tenenbaum et al. [2017]. Although social cues are often beneficial, attending to social stimuli can be detrimental if they are not aligned with the linguistic input most relevant for word learning. Greater attention to a speaker's mouth has been shown to be associated with better language abilities in children with ASD and typical development [Tenenbaum, Amso, Abar, & Sheinkopf, 2014]—potentially because attention to the mouth is beneficial for learning phonology [e.g., Lewkowicz & Hansen-Tift, 2012]. Given this evidence, Tenenbaum et al. hypothesized that directing children to focus on a speaker's mouth would also support the learning of new words. Contrary to expectations, children with ASD failed to learn new words when an adult pointed at her mouth while labeling a novel object. The authors proposed that, "pointing to the mouth of the speaker distracted the participants from the target object and resulted in failure to connect the object with the target label" [Tenenbaum et al., 2017, p. 12]—a clear example of how auditory–visual misalignment can prevent children from successfully learning new words.

Domain-General Attention

Although differences in social attention are a defining feature of ASD, differences in domain-general attention have also been identified. Individuals with ASD show differences in what aspects of the environment capture their attention (i.e., attentional preferences) and how long their attention is held by particular things. Attention allocation in individuals with ASD is strongly driven by perceptual salience, regardless of the social nature of the stimuli [Pierce et al., 2016; Pierce, Conant, Hazin, Stoner, & Desmond, 2011]. Individuals with ASD show an increased interest in geometric patterns [Amso, Haas, Tenenbaum, Markant, & Sheinkopf, 2014] and in certain objects, such as trains [Sasson, Elison, Turner-Brown, Dichter, & Bodfish, 2011; Sasson, Turner-Brown, Holtzclaw, Lam, & Bodfish, 2008; Unruh et al., 2016]. Eye-tracking studies [Sasson et al., 2008, 2011] have shown that children with ASD explore fewer stimuli in their environment (patterns of perseveration and circumscribed interest), and the things they look at, they explore in a more detail-oriented way. Once they are looking at something, they may have difficulty disengaging their attention—either taking longer to look away or not looking away at all [Elsabbagh et al., 2009, 2013; Landry & Bryson, 2004; Sacrey et al., 2014].

As with social attention, differences in domain-general attention can produce auditory–visual misalignment. Children will receive misaligned auditory and visual information if they look at something that interests them rather than something an adult is talking about, or if they fail to disengage attention in a timely manner. Even seemingly brief delays of a few hundred milliseconds may create misalignment at key moments in time. Because actions can be fleeting, action verbs may be particularly affected—a lag in shifting attention may mean that the visual stimuli that were relevant for learning are no longer aligned with incoming auditory information by the time the child shifts attention. Other types of words may be relatively less affected by difficulties in disengaging attention. For example, sound effects may frequently be paired with movement (e.g., when a parent moves a toy dog around while making barking sounds). This auditory–visual synchrony may be more likely to capture attention [Bahrick & Lickliter, 2000; Bahrick, Lickliter, & Flom, 2004], thereby increasing the likelihood that the child will look at the relevant visual information. Determining how different types of words (e.g., prepositions, adjectives) are relatively more or less affected by auditory–visual misalignment in children with ASD is an important area for future research.

The relationship between domain-general attentional differences and vocabulary development in children with ASD has received relatively limited empirical consideration. However, there is growing correlational evidence that domain-general attentional differences may produce auditory–visual misalignment that is detrimental for vocabulary development. Amso et al. [2014] found that children with ASD whose visual attention was driven most strongly by perceptual salience showed particularly severe impairments in language comprehension. In other work, researchers found that toddlers with ASD who preferred to look at geometric images (rather than social images) also had weaker language skills [Pierce et al., 2016]. Venker [2017] found that poorer visual disengagement was associated with poorer recognition of familiar words. Although the previous studies did not examine the impact of autism symptomology, Bavin et al. [2014] reported that in a word recognition task children with more severe autism symptoms were more likely to experience a mismatch between auditory and visual information. Bavin et al. [2014] suggested that, "Children who are slower or who are less likely to match auditory and visual information are at risk for misunderstanding the language addressed to them. Since linguistic input is so rapid, if children are slow even by a few hundred milliseconds in processing lexical items, the effect will accumulate over the course of a sentence…" (pp. 692–693). Though a full discussion is beyond the scope of this commentary, it is also possible that slowed auditory processing speed or other differences in auditory attention in individuals with ASD could lead to auditory–visual misalignment, even if auditory and visual input appear to be aligned from the perspective of the person providing linguistic input.

Increasing evidence of a link between domain-general attention and language in children with ASD has led some research groups to suggest that differences in

domain-general attention may negatively affect language development [Amso et al., 2014; Keehn et al., 2013; Pierce et al., 2016]. Although this work has considerably advanced the way we think about attention and language in individuals with ASD, the proposals made thus far have been relatively broad—perhaps due in part to the fact that these studies have primarily used omnibus measures of language ability that cover multiple domains (e.g., vocabulary, grammar, early literacy). The auditory–visual misalignment perspective proposed here builds on these previous proposals by identifying a specific aspect of language likely to be disrupted—vocabulary development—and a mechanism through which this disruption may take place—misalignment of auditory–visual statistics.

Conclusion

As described by Smith et al. [2014], "…the data available to any statistical learning machinery are not the data in the real world, but only a subset of that data that makes contact with the…learning system" (p. 254). Thus, children's attentional focus does more than simply limit what a child learns from the statistics in the environment that are relevant for early word learning—attention *determines* what those statistics are going to be. This is important to consider because the auditory–visual co-occurrences perceived by the child define what regularities are extracted from the input (i.e., what is learned). Infants and children, including those with ASD [Haebig et al., 2017; Mayo & Eigsti, 2012; Venker, in press], are sensitive to environmental statistics through mere exposure, even when they are not instructed to learn [Aslin, 2017; Erickson & Thiessen, 2015; Saffran, Aslin, & Newport, 1996; Vouloumanos & Werker, 2009]. If looking at the right thing at the right time facilitates vocabulary development, then looking at the wrong thing is likely to be quite problematic.

In this commentary, we have presented evidence that because of a host of attentional differences in children with ASD, the linguistic input these children hear is less likely to be related to what they see at any given time—a phenomenon we refer to as auditory–visual misalignment. We have also presented evidence that because auditory–visual alignment disrupts the statistics available for word learning, it is likely to disrupt vocabulary development in children with ASD. Theoretically, this perspective of auditory–visual misalignment advances our understanding of how attentional differences impact vocabulary development in children with ASD in three primary ways.

First, by adopting the point of view of the child—the auditory–visual statistics a child is taking in—we provide a framework that brings together research on social and domain-general visual attention differences in children with ASD. This perspective emphasizes not the characteristics of visual stimuli in isolation (e.g., whether a child is looking at a person or an object), but whether the visual stimuli a child is looking at are relevant to the linguistic input. That is, attention to an irrelevant stimulus produces misalignment and detracts from a learning opportunity, regardless of whether that stimulus is social or non-social [see Tenenbaum et al., 2017].

Second, the auditory–visual misalignment perspective moves current thinking beyond how misalignment disrupts vocabulary development in the moment, and considers the likely consequences of misalignment over developmental time [Kucker et al., 2015]. Typically developing infants experience transient phases of inflexible visual attention driven by perceptual salience [Hollich et al., 2000; Johnson, Posner, & Rothbart, 1991; Pruden et al., 2006; Smith et al., 2014]. In contrast, children with ASD experience these characteristics of visual attention for a protracted time, perhaps through adulthood [Kawakubo et al., 2007], likely resulting in cumulative effects on vocabulary development. Even subtle misalignment accruing over time would create inconsistencies and ambiguities in auditory–visual statistics, altering and/or slowing the development of robust lexical representations [also see Baron-Cohen et al., 1997]. In this way, misalignment may help to explain why even children with ASD who eventually develop age-appropriate or precocious language abilities show early language delays, despite strong processing and learning skills.

Finally, considering auditory–visual misalignment may assist in identifying active ingredients of existing language interventions or in developing new interventions that deliver high quality, aligned input. For example, this perspective provides a framework for understanding how parental verbal responsiveness might facilitate vocabulary development over time [also see Tamis-Lemonda et al., 2014]. McDuffie and Yoder [2010] found that the frequency of parent comments related to the child's focus of attention uniquely predicted children's later spoken vocabulary, controlling for child engagement with objects. Thus, children with ASD demonstrated better vocabulary outcomes when their parents used a strategy that facilitated auditory–visual alignment. In describing the value of follow-in comments for facilitating vocabulary development, McDuffie and Yoder [2010] stated that, "Such utterances would make the associative pairing between label and referent more explicit and would take advantage of temporal contiguity as a passive support to facilitate accuracy and efficient word learning" (p. 11).

Of course, the perspective described here is only one of many potential explanations of vocabulary delays in children with ASD; many other factors are likely to play a role. In particular, social communication remains an important part of vocabulary development. Children

with ASD have difficulty monitoring and assessing social intent, which can make word learning more difficult [Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff, & Tager-Flusberg, 2007; Preissler & Carey, 2005] and may lead to inappropriate generalization of word meanings [McGregor & Bean, 2012].

In our view, the auditory–visual misalignment perspective identifies several important avenues for future research. The explanatory value of auditory–visual misalignment should be tested by prospective, longitudinal studies of young children with or at risk for ASD that measure children's language and attention skills—as well as their language-learning environments—at multiple points over early development. All else equal, this perspective would predict that children with ASD who experience higher rates of auditory–visual alignment would develop stronger vocabulary skills. Furthermore, individual differences in the amount of auditory–visual misalignment children experience may impact vocabulary development. Additional work is needed to determine the extent to which auditory–visual misalignment may produce incorrect learning [as seen in the study by Baron-Cohen et al., 1997] versus slow learning or the absence of correct learning. Such work would also shed light on the learning mechanisms that underlie vocabulary development—an issue of theoretical interest in typical development [Smith & Yu, 2008; Trueswell, Medina, Hafri, & Gleitman, 2013] as well as atypical development. Future research will also be required to determine how manipulating auditory–visual alignment changes word learning in children with ASD and whether the effects of auditory–visual misalignment are unique to ASD or shared with other neurodevelopmental disorders or sources of language impairment.

Acknowledgment

This work was supported in part by NIH grant R21 DC016102 (Venker, PI).

References

Amso, D., Haas, S., Tenenbaum, E., Markant, J., & Sheinkopf, S. J. (2014). Bottom-up attention orienting in young children with autism. Journal of Autism and Developmental Disorders, 44, 664–673. https://doi.org/10.1007/s10803-013-1925-5.

Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. Autism Research, 9, 810–828. https://doi.org/10.1002/aur.1590.

Aslin, R. N. (2017). Statistical learning: A powerful mechanism that operates by mere exposure. Wiley Interdisciplinary Reviews: Cognitive Science, 8(8), 1–7.

Axelsson, E. L., Churchley, K., & Horst, J. S. (2012). The right thing at the right time: Why ostensive naming facilitates word learning. Frontiers in Psychology, 3, 1–8. https://doi.org/10.3389/fpsyg.2012.00088.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. Developmental Psychology, 36, 190–201. https://doi.org/10.1037/0012-1649.36.2.190.

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in fnfancy. Current Directions in Psychological Science, 13, 99–102. https://doi.org/10.1111/j.0963-7214.2004.00283.x.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. Child Development, 55, 1278–1289.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? Child Development, 68, 48–57.

Bavin, E. L., Kidd, E., Prendergast, L., Baker, E., Dissanayake, C., & Prior, M. (2014). Severity of autism is related to children's language processing. Autism Research, 7, 687–694. https://doi.org/10.1002/aur.1410.

Bono, M. A., Daley, T., & Sigman, M. (2004). Relations among joint attention, amount of intervention and language gain in autism. Journal of Autism and Developmental Disorders, 34, 495–505. https://doi.org/10.1007/s10803-004-2545-x.

Bottema-Beutel, K. (2016). Associations between joint attention and language in autism spectrum disorder and typical development: A systematic review and meta-regression analysis. Autism Research, 9, 1021–1035. https://doi.org/10.1002/aur.1624.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. Developmental Science, 8, 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A., & Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. International Journal of Language & Communication Disorders, 38, 265–285. https://doi.org/10.1080/136820310000104830.

Charman, T., Howlin, P., Berry, B., & Prince, E. (2004). Measuring developmental progress of children with autism spectrum disorder on school entry using parent report. Autism, 8, 89–100.

Ellis Weismer, S., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. Journal of Autism and Developmental Disorders, 40, 1259–1273. https://doi.org/10.1007/s10803-010-0983-1.

Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. Biological Pshychiatry, 74, 189–194.

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L. Csibra, G., Baron-Cohen, S., & Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: Disengagement and facilitation. Journal of Child Psychology and Psychiatry, 50, 637–642.

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning

account of language acquisition. Developmental Review, 37, 66–108. https://doi.org/10.1016/j.dr.2015.05.002.

Foss-Feig, J. H., Schauder, K. B., Key, A. P., Wallace, M. T., & Stone, W. L. (2017). Audition-specific temporal processing deficits associated with language function in children with autism spectrum disorder. Autism Research: Official Journal of the International Society for Autism Research, 10, 1845–1856. https://doi.org/10.1002/aur.1820.

Haebig, E., Saffran, J. R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. Journal of Child Psychology and Psychiatry, 58, 1251–1263. https://doi.org/10.1111/jcpp.12734.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., & Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. Monographs of the Society for Research in Child Development, 65, 1–135.

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. Journal of Cognitive Neuroscience, 3, 335–344. https://doi.org/10.1162/jocn.1991.3.4.335.

Kasari, C., Gulsrud, A., Freeman, S., Paparella, T., & Hellemann, G. (2012). Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. Journal of the American Academy of Child & Adolescent Psychiatry, 51, 487–495. https://doi.org/10.1016/J.JAAC.2012.02.019.

Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: Randomized comparison of joint attention and play interventions. Journal of Consulting and Clinical Psychology, 76, 125–137 Retrieved from http://psycnet.apa.org/fulltext/2008-00950-015.html.

Kasari, C., Sigman, M., Mundy, P., & Yirmiya, N. (1990). Affective sharing in the context of joint attention interactions of normal, autistic, and mentally retarded children. Journal of Autism and Developmental Disorders, 20, 87–100. https://doi.org/10.1007/BF02206859.

Kawakubo, Y., Kasai, K., Okazaki, S., Hosokawa-Kakurai, M., Watanabe, K., Kuwabara, H., … Maekawa, H. (2007). Electrophysiological abnormalities of spatial attention in adults with autism during the gap overlap task. Clinical Neurophysiology, 118, 1464–1471. https://doi.org/10.1016/J.CLINPH.2007.04.015.

Keehn, B., Muller, R.-A., & Townsend, J. (2013). Atypical attentional networks and the emergence of autism. Neuroscience & Biobehavioral Reviews, 37, 164–183. https://doi.org/10.1016/j.neubiorev.2012.11.014.Atypical.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2015). Slowing down fast mapping: Redefining the dynamics of word learning. Child Development Perspectives, 9, 74–78. https://doi.org/10.1111/cdep.12110.

Landry, R., & Bryson, S. E. (2004). Impaired disengagement of attention in young children with autism. Journal of Child Psychology and Psychiatry, 45, 1115–1122.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. In Proceedings of the National Academy of Sciences of The United States of America (Vol. 109, pp. 1431–1436). https://doi.org/10.1073/pnas.1114783109.

Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. Journal of Autism and Developmental Disorders, 38, 1426–1438. https://doi.org/10.1007/s10803-007-0510-1.

Mayo, J., & Eigsti, I.-M. (2012). Brief report: A comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. Journal of Autism and Developmental Disorders, 42, 2476–2485. https://doi.org/10.1007/s10803-012-1493-0.

McDuffie, A., & Yoder, P. (2010). Types of parent verbal responsiveness that predict language in young children with autism spectrum disorder. Journal of Speech, Language, and Hearing Research, 53, 1026–1039. https://doi.org/10.1044/1092-4388(2009/09-0023.

McGregor, K. K., & Bean, A. (2012). How children with autism extend new words. Journal of Speech, Language, and Hearing Research, 55, 70–84. https://doi.org/10.1044/1092-4388(2011/11-0024)b.

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. Infant Behavior and Development, 21, 469–482. https://doi.org/10.1016/S0163-6383(98)90020-0.

Mundy, P., & Jarrold, W. (2010). Infant joint attention, neural networks and social cognition. Neural Networks, 23, 985–997. https://doi.org/10.1016/J.NEUNET.2010.08.009.

Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. Journal of Autism and Developmental Disorders, 20, 115–128. https://doi.org/10.1007/BF02206861.

Naigles, L. R., Kelty, E., Jaffery, R., & Fein, D. (2011). Abstractness and continuity in the syntactic development of young children with autism. Autism Research, 4, 422–437. https://doi.org/10.1002/aur.223.

Parish-Morris, J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., & Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intention in word learning. Child Development, 78, 1265–1287. https://doi.org/10.1111/j.1467-8624.2007.01065.x.

Paul, R., Chawarska, K., Cicchetti, D., & Volkmar, F. (2008). Language outcomes of toddlers with autism spectrum disorders: A two year follow-up. Autism Research, 1, 97–107. https://doi.org/10.1002/aur.12.

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. Psychonomic Bulletin & Review, 21, 178–185. https://doi.org/10.3758/s13423-013-0466-4.

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. Archives of General Psychiatry, 68(1), 101–109. https://doi.org/10.1001/archgenpsychiatry.2010.113.

Pierce, K., Marinero, S., Hazin, R., McKenna, B., Barnes, C. C., & Malige, A. (2016). Eye-tracking reveals abnormal visual preference for geometric images as an early biomarker of an ASD subtype associated with increased symptom severity. Biological Psychiatry, 79, 657–666. https://doi.org/10.1016/j.biopsych.2015.03.032.

Pomper, R., & Saffran, J. R. (2018). Familiar object salience affects novel word learning. Child Development, 1–17.

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. Cognition, 9(1), B13–B23. https://doi.org/10.1016/j.cognition.2005.01.008.

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. Child Development, 77, 266–280. https://doi.org/10.1111/j.1467-8624.2006.00869.x.

Rovee-Collier, C. (1995). Time windows in cognitive development. Developmental Psychology, 31, 147–169.

Sacrey, L.-A. R., Armstrong, V. L., Bryson, S. E., & Zwaigenbaum, L. (2014). Impairments to visual disengagement in autism spectrum disorder: A review of experimental studies from infancy to adulthood. Neuroscience & Biobehavioral Reviews, 47, 559–577. https://doi.org/10.1016/j.neubiorev.2014.10.011.

Saffran, J. R., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. Science, 274, 1926–1928.

Sasson, N. J., Elison, J. T., Turner-Brown, L. M., Dichter, G. S., & Bodfish, J. W. (2011). Brief report: Circumscribed attention in young children with autism. Journal of Autism and Developmental Disorders, 41, 242–247. https://doi.org/10.1007/s10803-010-1038-3.

Sasson, N. J., Turner-Brown, L. M., Holtzclaw, T. N., Lam, K. S. L., & Bodfish, J. W. (2008). Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. Autism Research, 1, 31–42. https://doi.org/10.1002/aur.4.

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. Trends in Cognitive Sciences, 18, 251–258. https://doi.org/10.1016/j.tics.2014.02.007.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. Cognition, 106, 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010.

Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. Child Development, 78, 542–557. https://doi.org/10.1111/j.1467-8624.2007.01022.x.

Tamis-Lemonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? Current Directions in Psychological Science, 23, 121–126. https://doi.org/10.1177/0963721414522813.

Tenenbaum, E. J., Amso, D., Abar, B., & Sheinkopf, S. J. (2014). Attention and word learning in autistic, language delayed and typically developing children. Frontiers in Psychology, 5, 490. https://doi.org/10.3389/fpsyg.2014.00490.

Tenenbaum, E. J., Amso, D., Righi, G., & Sheinkopf, S. J. (2017). Attempting to "Increase Intake from the Input": Attention and word learning in children with autism. Journal of Autism and Developmental Disorders, 47, 1791–1805. https://doi.org/10.1007/s10803-017-3098-0.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. Child Development, 57, 1454–1463.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. Cognitive Psychology, 66, 126–156. https://doi.org/10.1016/j.cogpsych.2012.10.001.

Unruh, K. E., Sasson, N. J., Shafer, R. L., Whitten, A., Miller, S. J., Turner-Brown, L., & Bodfish, J. W. (2016). Social orienting and attention is influenced by the presence of competing nonsocial information in adolescents with autism. Frontiers in Neuroscience, 10, 1–12. https://doi.org/10.3389/fnins.2016.00586.

Venker, C. E. (in press). Cross-situational and ostensive word learning in children with and without autism spectrum disorder. Cognition.

Venker, C. E. (2017). Spoken word recognition in children with autism spectrum disorder: The role of visual disengagement. Autism, 2, 821–829.

Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. Developmental Psychology, 45, 1611–1617. https://doi.org/10.1037/a0016134.