

**Deliberate synchronization of speech and gesture:
Effects of neurodiversity and development**
Inge-Marie Eigsti¹ and Wim Pouw²

¹Department of Psychological Sciences, University of Connecticut; 406 Babbidge Road; Storrs, CT 06269; USA

²Donders Center for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

Corresponding author: Inge-Marie Eigsti, Ph.D., University of Connecticut; Department of Psychological Sciences; 406 Babbidge Road; Storrs, CT 06269; USA. Email: inge-marie.eigsti@uconn.edu; Telephone: 860-486-6021

Keywords: autism spectrum disorder, beat gesture, perception-action coordination, synchrony

Acknowledgments. This research was supported by NIMH-1R01MH112687-01A1 to Eigsti and Fein (Co-PIs) and by NIDCD T32DC017703 to Eigsti and Myers (Co-PIs). The authors thank He and colleagues for sharing their Praat code.

Open data statement: The processed datasets (i.e., timing data) are available on Open Science Framework (link: <https://osf.io/kcr8x/> DOI 10.17605/OSF.IO/KCR8X). The raw audio-visual data are stored by the first author, but contain sensitive data that cannot be openly shared. Since the project was stalled for several years, during which the second author changed institutions, the project files that contained the processing scripts, statistical analysis scripts, raw time series data, and a copy of the audiovisual recordings was deleted. However, the analyses can be reproduced based on descriptions in the paper, and the authors are happy to share relevant information upon inquiry.

Running head: Speech-movement synchrony in autism

Author contributions: IME conceived of the study, and lead writing of the introduction, method (design and experiment), and discussion the of manuscript, with critical revisions by WP. WP conceived and performed the analyses and led the writing of the results and method (analyses), with critical revisions by IME.

Conflict of Interest: The authors declare that they have no conflict of interest.

Submitted to Language and Cognition, Special issue: “Multimodal prosody: Speech and gesture in interaction”

Abstract

The production of speech and gesture is exquisitely temporally coordinated. In autistic individuals, speech-gesture synchrony during spontaneous discourse is disrupted. To evaluate whether this asynchrony reflects motor coordination versus language production processes, the current study examined *deliberately* performed hand movements during speech in youth with autism spectrum disorder (ASD) compared to neurotypical youth. Neurotypical adult performance provided a mature baseline. Participants read aloud rhythmic nursery rhymes, while producing a beat-like hand movement. An automated pixel-change video measure identified kinematic peaks; using smoothed acoustic envelope analyses, we identified peaks in speech. Results indicated few diagnostic group differences in explicit speech-movement coordination, although adolescent performance differed from adults. Adults demonstrated higher tempo and greater rhythmicity in their coordination; this group difference suggests that the method is sufficiently subtle to reveal individual differences and that this form of complex coordination undergoes ongoing maturation beyond adolescence. The sample is small, and thus results are necessarily preliminary. In the context of prior speech-gesture coordination studies, these findings of intact synchrony are consistent with the hypothesis that it is the demands of discourse planning, rather than motor coordination, that have led to prior findings of asynchrony during spontaneous speech; this possibility awaits future research.

Introduction

Spontaneous co-speech gestures play a central role in language acquisition (Brentari & Goldin-Meadow, 2017), language processing (Alibali et al., 2000; McNeill, 1985, 1992), and learning (Cook et al., 2008; Goldin-Meadow et al., 2009; Goldin-Meadow & Wagner, 2005). However, the gestures produced by individuals with autism spectrum disorder (ASD) show striking developmental and maturational differences from the gestures produced by their neurotypical peers. Fluent autistic talkers display an asynchrony of speech and gesture coordination, with gestures both preceding and lagging the associated speech (de Marchena & Eigsti, 2010; Morett et al., 2016). Similar findings have been reported for gesture/gaze asynchrony (Bloch et al., 2022) and visual/motor asynchrony (Nebel et al., 2016). Speech-gesture synchrony impacts listener comprehension, such that narratives characterized by more asynchronous gestures are rated by naïve listeners as more difficult to follow (de Marchena & Eigsti, 2010). That is, autistic narrators produced more iconic gestures that were poorly synchronized with the *semantically* relevant speech (e.g., a “throwing” gesture that is temporally displaced from the phrase “and he threw it”); furthermore, the degree of temporal asynchrony was strongly associated with naïve ratings of how easy a story was to understand. Motor coordination itself is often impaired in autism (indeed, some have suggested it be added as a diagnostic criterion of autism; Bhat, 2021). As such, this raises an important question: is autistic speech-gesture asynchrony due to impairments in motor coordination, or is it a reflection of impairments in higher-order language processes such as discourse planning? The current study was designed to investigate the deliberate coordination of hand movements with speech, in autistic individuals, to help rule out the impact of motor processes in gesture asynchrony in autism.

The study focuses on beat-like rather than representational qualities of gesture as related to development and neurodiversity. In typical adults, pure beat gestures are structured, pulse-like movements (typically, a simple up-and-down “biphasic” motion) that couple with the prosody of speech, conveying pragmatic information with limited semantic or referential content (Pouw & Dixon, 2019; Wagner et al., 2014). However, some semantically laden iconic gestures may also have a beat-like speech-synchronized quality (Pouw & Dixon, 2019; Rohrer et al., 2023). Beat gestures are closely synchronized with acoustically prominent accented syllables (Krahmer & Swerts, 2007; Wang & Chu, 2013). Studies of the development of beat gestures suggest such gestures contribute to listener comprehension; for example, children who actively move their hands while telling a story produced more prosodically fluent stories with stronger narrative structure compared to children who are not encouraged to gesture (Vilà-Giménez & Prieto, 2020). Most studies of gesture production in development and in autism have not specifically examined beat gestures (see below), but what we do know suggests some alternative developmental patterns in this domain. We review here the literature on gesture production in autism, speech-gesture synchrony and relevant studies of synchrony and related phenomena in autism, and studies of motor control and coordination in autism.

Gesture development in autism. Mounting evidence indicates that the development of gestures is altered in autism¹. Clinically, impairments in gesture are asserted on gold-standard ASD diagnostic measures and screeners such as the Autism Diagnostic Observation Schedule (Lord et al., 2012), the Autism Diagnostic Interview (ADI; Lord et al., 1994), and the Modified

¹ Regarding terminology (how to talk about autism), preferences of the autism community encompass both person-first (e.g., “adolescents with autism”) and identity-first (e.g., “autistic adolescents”) approaches (Keating et al. (2023) Respecting these preferences, this manuscript utilizes both terms.

Checklist for Autism in Toddlers (M-CHAT; Robins et al., 2001); on these measures, absence or infrequent gestures is thought to be symptomatic. Deictic gestures early in development in autism have been well-studied; results suggest that early *declarative deictics* (i.e., pointing to share attention), are reduced in frequency, whereas *instrumental deictics* (i.e., pointing to request), are less affected, which suggests an impairment in the communicative rather than the motoric domain (Bono et al., 2004; Loveland & Landry, 1986a; Mundy et al., 1990; Mundy et al., 1986). The expression and comprehension of deictics is reduced (Mundy et al., 1986) and delayed (Camaioni et al., 1997) in autism, and these delays are tied to delays in language acquisition (Loveland & Landry, 1986b; Mundy et al., 1990). Some studies report reduced rates of gesture in children with ASD (Bartak et al., 1975, 1977; Bono et al., 2004), and most studies report delays in the onset of gesture production (Charman et al., 2003; Luyster et al., 2007). Children with ASD produce a reduced variety of gestures (Colgan et al., 2006; Wetherby & Prutting, 1984), and their gestures are less likely to occur in combination with both vocalizations and eye contact (Murillo et al., 2021). Gesture production in autism is a strong predictor of later communication skills (Taverna et al., 2021), as reviewed by Ramos-Cabo et al. (2019). Speech and gesture development is longitudinally closely connected in autistic toddlers as is true in typical development (Dimitrova & Özçalışkan, 2022; Ingersoll & Lalonde, 2010; Özçalışkan et al., 2017; Tager-Flusberg et al., 1990). Overall, the literature suggests significant delays in early gesture development in autism, particularly for more semantically complex and more social gestures, and suggests that these early delays correlate with later language attainment.

Gestures in verbally fluent autistic individuals. While most of the autism gesture literature has focused on the period of early language acquisition, a growing empirical literature describes gesture production impairments in verbally fluent children and adults. Many studies examine spontaneous gesture production during narrative tasks. Some studies report reduced gesture rates in autism (Silverman et al., 2017), while others find no group differences in gesture frequency after accounting for number of utterances (Attwood et al., 1988; de Marchena & Eigsti, 2010; Garcia-Perez et al., 2007; Tantam et al., 1993; Wong & So, 2018). A study of verbally fluent autistic and non-autistic school-age children used an experimental task eliciting gestures, and found that the autism group was less likely to gesture to specific spatial locations to refer to non-present events or objects (So et al., 2015). A study of conversation in verbally fluent autistic adults and their neurotypical peers reported group differences in both semantic/pragmatic and motoric features of spontaneously produced co-speech gestures (de Marchena et al., 2019). The autistic adults were more likely to use gesture to facilitate turn-taking in conversation, and produced more unimanual than bimanual gestures. De Marchena and Eigsti (2014) found that adolescents with ASD produced *fewer* gestures while telling a story to a listener, but produced *more* deictic gestures when completing an individual, non-communicative executive function task, compared to an age- and IQ-matched non-autistic group, suggesting that the gestures were used more to regulate one's own processing rather than for communicating with a listener. Relatedly, Morett and colleagues (2016) coded gestures produced in the presence or absence of a visible listener, and found that, for the non-autistic comparison group, communicative quality and gesture frequency increased in the presence of a visible listener; there was no such increase in the autism group, suggesting that social impairments contribute to gesture production differences.

Speech-gesture synchrony. Gesture and speech have a tightly synchronized timecourse that connects to semantic and pragmatic features (Levelt et al., 1985; Nobe, 2010) and prosodic features (Wagner et al., 2014); they form an integrated system through which we produce and

comprehend meaning (Kelly et al., 2010). In addition to coupling and coordination on a conceptual (e.g., at a conceptual, semantic level), synchrony (coupling) may also reflect purely motoric processes required to coordinate motor planning for speech with that involved in gesture production. For example, when speaking mono-syllable utterances while simultaneously tapping a finger, adult speakers tend to transfer an emphasis from gesture to speech, and from speech to gesture (Parrell et al., 2014). Specifically, during more emphatic tapping, mouth aperture increases, and when producing more emphatic speech, tapping motions are larger. Similarly, in a recent study it was found that more beat-like (decelerative) gestures during counting-out-rhymes were associated with more emphasized speech (Kadavá et al., August 7-8, 2023). Such gesture-speech coupling has been argued to be supported by biomechanical interactions of the upper limbs and the respiratory-vocal system (for a review, see Pouw & Fuchs, 2022). In this account, gestures play a direct physical role in the control of vocalization; that is, the biomechanics of motor control push vocalizations and gestures into synchronized activity. Consistent with this hypothesis, results indicate that in the context of severe impairments in speech production (e.g., aphasia), aspects of gesture-speech synchrony such as coupling between gesture-kinematics and vocal-acoustics are maintained to some degree (Jenkins & Pouw, 2023; Pouw et al., 2022).

In autism, several studies of semantic speech-gesture synchrony and movement timing report clear group differences. An older study reported a significant reduction in the co-occurrence of gestures with vocalization in autism (Tantam et al., 1993); this finding has been conceptually replicated by more recent studies studying the timing between gesture strokes and their lexical affiliates (de Marchena & Eigsti, 2010; Morett et al., 2016). In the de Marchena study, the degree of asynchrony of speech and gestures predicted ratings of communicative quality and autism symptom severity, indicating that speech-gesture synchrony was correlated with comprehension in listeners who saw videos of story narrations and were asked to rate the clarity of the stories (see also Habets et al., 2010).

Autistic gestures appear atypical in dimensions beyond synchrony with speech, specifically in kinematics, or motion velocity. These differences may be early-emerging; a retrospective review of home videos of infants later diagnosed with autism suggested timing differences in early bimanual, but not unimanual, repetitive movements (Purpura et al., 2017). Using a motion-tracking system, one study reported significant kinematic differences in the number of meaningful holds between movements in the autism group (Trujillo et al., 2021). Another motion-tracking study examined the production of nonverbal signals toward a target that was invisible to a communication partner, and found greater delays (reduced temporal coherence) of nonverbal gestures in the adult autistic group (Bloch et al., 2022). A study of the kinematics of simple movements in autistic adults found group differences in jerk, acceleration, and velocity, suggesting atypical movement quality (Cook et al., 2013). Similarly, a study of 9-14 year old autistic and non-autistic children revealed an atypical kinematic profile that was correlated with differences in the perception of biological motion and with measures of autism characteristics (Butera et al., 2017). A study of the dynamics of movement in autistic youth found that while grasping activity was typical during a non-social activity, the autism group displayed faster transitioning of grasping activity due to object size changes (reduced hysteresis) when a social component was added to the task as compared to a control group (Amaral et al., 2017). Although a study of interpersonal coordination of whole-body movements during conversation found no group differences (Romero et al., 2018), most research to date examining subtle motor dynamics reports significant differences in autism.

Motor impairments in autism. As described, studies in autism have revealed generally atypical gesture development, atypical speech-gesture coordination, and atypical motor kinematics. Of course, these findings could all be due to difficulties with motor coordination. Numerous studies suggest that autism is characterized by important developmental differences in motor coordination (Macneil & Mostofsky, 2012; McAuliffe et al., 2017; Mostofsky & Ewen, 2011; Nebel et al., 2016). Motor impairments in movement preparation (Rinehart et al., 2001), movement coordination (McAuliffe et al., 2017; Vilensky et al., 1981), and differences in reaching and grasping movements (Glazebrook et al., 2009), all serve as important predictors of long-term outcomes in core ASD domains such as social and communication skills (Stevens et al., 2000). A study of 95 autistic toddlers followed longitudinally found that, somewhat surprisingly, early motor skills as assessed in infancy were a *stronger predictor* of later social and communicative outcomes, than the severity of autism symptomatology (Sutera et al., 2007); see also (Eigsti, 2024). This evidence supports the importance of motor skills in autism for later outcomes. While motor coordination is not currently part of the diagnostic criteria for ASD, it may be central to the phenotype. Multiple studies report significant differences and deficits in autistic samples, compared to non-autistic individuals (Cattaneo et al., 2007; Dziuk et al., 2007; Fabbri-Destro et al., 2009; Hughes, 1996; Jansiewicz et al., 2006; Macneil & Mostofsky, 2012; Scharoun & Bryden, 2016). There are inconsistencies; for example, some studies report age-appropriate motor adaptation to changes in the environment (Gidley Larson et al., 2008) and no impairment in simple (one-step) action imitation tasks (Hamilton et al., 2007; van Swieten et al., 2010). However, given the prevalence of motor delays and impairments in ASD (Bhat et al., 2011), and the relevance of such difficulties to diagnostic classification (Harrison et al., 2021), some have proposed that motor deficits be incorporated into diagnostic criteria (Bhat, 2021). As such, speech-gesture asynchrony in autism could be primarily due to impairments in motor skills, rather than impairments in language production.

Rhythmic speech in autism. Speech-gesture coordination requires precise timing of speech production. The production of rhythmic speech in autism has not been well studied, but research to date suggests that phonological awareness (central to speech and reading) was correlated with musical beat perception in autistic children (Rimmer et al., 2023). A larger study of 78 autistic and 84 non-autistic adults found intact perception of rhythmic information, but a reduction in rhythmic entrainment in the autism group (Cannon et al., 2023). Studies of rhythmic production are limited, but suggest autistic impairments in rhythmic prosodic production and articulatory timing (Lau et al., 2023).

In summary, research to date suggests that gesture development and gesture production in autism is atypical. Autistic people have significant asynchrony of speech-gesture coordination, and this asynchrony has a significant negative impact on listener perception of meaning (de Marchena & Eigsti, 2010). While many have interpreted this as related to broader impairments in language and communication, the literature reviewed above also suggests autism-specific difficulties with subtle motor dynamics as well as broader deficits in motor coordination of both hand movements and of the motor aspects of speech planning.

The current study. Motor coordination is critical for producing gestures that are well synchronized with speech. It is an open question whether speech-gesture asynchrony in autism is primarily a motor control issue, or arises due to broader communication difficulties such as discourse planning (of course, these possibilities are not mutually exclusive and both may influence performance simultaneously). To address this question, an exploratory study probed for differences in the deliberate coordination of speech with beat-like motor movements,

comparing autistic and non-autistic individuals matched on age, gender, and nonverbal cognitive abilities. The approach used sensorimotor synchronization methods (Repp, 2005) and results were analyzed using an approach developed in multimodal studies of gesture-speech synchrony (Pouw et al., 2020). The study focused on coordination in verbally fluent adolescents ages 12-17. This age range was chosen because the process of language acquisition is largely complete at this point. However, adolescents undergo rapid changes in body size and shape, including changes in limb length that are relevant to gesture production, and continued improvements in motor coordination (Kemper et al., 2015) as well as extensive changes in neural structure and organization (Fuhrmann et al., 2015). The transition from adolescence into young adulthood may be marked by meaningful changes in speech-gesture coordination; thus, this study included a non-autistic adult sample, to provide a mature stable baseline for comparison. This study utilized a deliberate, motoric task with reduced utterance planning demands, designed as an “analogue” to speech-gesture coordination during spontaneous speech. Asynchrony in this task was expected to reflect motor impairments, whereas intact synchrony would be consistent with the hypothesis that speech-gesture coordination deficits in autism instead reflect impairments in discourse planning or other linguistic processes.

Methods

Participants. This study included participants from three groups; see Table 1. Adolescents with ($n = 9$) and without ($n = 10$) autism diagnoses provided the primary comparison of interest. Adolescents were recruited through clinical contacts, local schools, and resource fairs, study flyers, and by word of mouth. Inclusion criteria for the autism group were a parent-reported diagnosis of ASD, which was validated by clinical evaluation in the study (see below). Exclusion criteria were major psychopathology that would preclude participation (schizophrenia or bipolar disorder), ascertained via parent report, and uncorrected hearing or vision impairments; in addition, non-autistic participants could not have a first-degree autistic relative or parent-reported history of significant developmental delay. In addition to the gesture task, participants completed a short battery of measures to capture cognitive and language functioning and executive processes, described below. All participants had nonverbal cognitive abilities in the average or high average range; see Table 1 for details. A convenience sample of non-autistic college students ages 18-22 years ($n = 11$) provided an adult performance baseline; these participants completed only the gesture-speech synchrony task. All participants spoke English as a primary or first language, though some participants were bilingual speakers; these data were not systematically recorded and thus were not considered in further analyses. Race and ethnicity data were not recorded and are thus not reported.

Table 1
Participant demographics

	Autism group <i>n</i> = 9	Non-autistic group; <i>n</i> = 10	χ^2/F	<i>p</i>	<i>d/w</i>
Age (years)	14.7 (2.0); 11-17	14.3 (1.4); 12-17	0.29	.60	3.23
Male: Female	7:2	8:2	0.01	.91	3.02
Full-scale IQ SS	101 (8); 85 - 109	104 (10); 100-114	0.45	.51	3.33
CELF Core Lang. SS	109 (9); 96-123	112 (12); 93-129	0.35	.56	3.28
Working Memory (Letter-Number Sequencing)	10.8 (3.5); 6 - 16	12.0 (3.8); 9 – 19	0.53	.48	0.33

Planning (Tower of London)	8.8 (2.1); 4-11	11.0 (1.8); 9-13	6.07	.03	1.12
ADOS Calibrated Severity Score	7.6 (1.8); 4-10	N/A			
Social Communication Questionnaire score	24.1 (5.0); 19-33	3.2 (2.5); 0-7	139.92	<.001	5.36

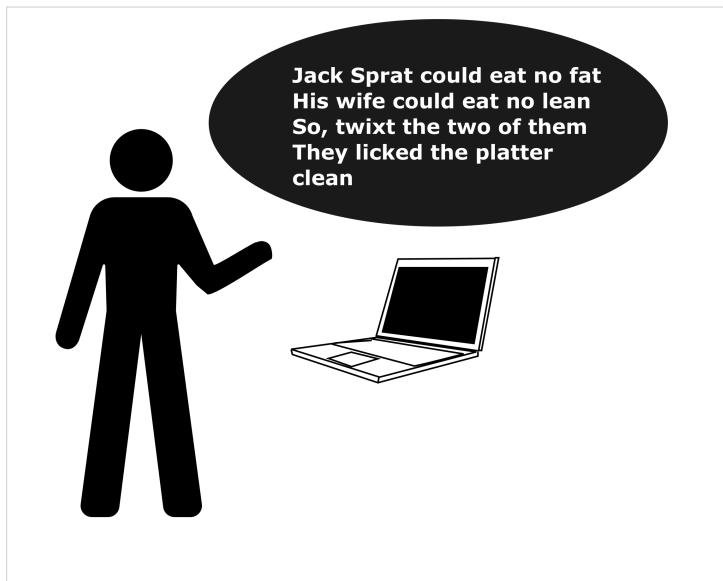
Notes: Data are presented as $M(SD)$; range. Standard scores (SS) = 100(15). Full-scale IQ = Wechsler Adult Scales of Intelligence (Wechsler, 2011) SS. CELF Core Lang. = Clinical Evaluation of Language Fundamentals, 4th Edition SS (CELF-4; Semel et al., 2003) . Working memory = Wechsler Intelligence Scales for Children – 4 (Wechsler, 2003) Letter-Number Sequencing SS. Planning = D-KEFS (Delis et al., 2001) Tower of London subscale scores $M(SD) = 10(3)$. ADOS = Autism Diagnostic Observation Schedule (Module 3; Lord et al., 2012); higher scores indicate greater symptomatology, and scores >4 are in the autism spectrum range. SCQ = Social Communication Questionnaire (Rutter et al., 2003); higher scores indicate greater symptomatology, and scores >15 are in the autism spectrum range. Adult participants ($n = 11$) did not complete clinical assessments, and are not included in this table.

Procedures

All study procedures were approved by the [redacted for review] Institutional Review Board. Participants were seen in a quiet room at their own home or in the lab, according to their preference. They recited six nursery rhymes in English while using their writing hand “as if you are hammering a hammer;” see Figure 1 for a schematic. They completed six trials in total; the six rhyming verses, which had three to six words per line, were *Jack Sprat Could Eat No Fat; Mary, Mary, Quite Contrary; Peter, Peter, Pumpkin Eater; Hickory, Dickory, Dock; Sing a Song of Sixpence*; and *There Was a Crooked Man*. The nursery rhymes were printed on a computer screen; participants were invited to read and review each rhyme until they felt comfortable with it. To decrease working memory demands, the text remained onscreen for the entire trial; most participants referred to the screen during the trial. They received no feedback on their performance; and indeed, none was needed, as all participants were able to comply with the task without training without further questions about the procedure. No trials were repeated. Most of the rhymes were at least somewhat familiar to most (but not all) participants. Participants were asked to stand while reciting the rhyme, to avoid interference from table tops or chair arms. Performance was recorded on digital video (25 frames per second), and audio was recorded with an external Shure PG42USB Cardioid condenser microphone. The motion peaks in the gestures and the acoustic peaks in the speech signal were identified and probed for coordination, as detailed below.

Figure 1.

Schematic of participant completing the explicit gesture task, moving the right arm up and down while reciting a rhyme stimulus viewed on a laptop computer.



In addition to the nursery rhyme task, the adolescents completed standardized assessments to evaluate the role of cognitive and diagnostic contributors to coordination. These assessments were the Wechsler Abbreviated Intelligence Scales (Wechsler, 2011), which is normed for ages six to 90 years, as a measure of Full-scale IQ; the Clinical Evaluation of Language Fundamentals, 4th Edition (CELF-4; Semel et al., 2003) Core Language composite score, as a measure of structural language abilities; the Letter-Number Sequencing subtest of the Wechsler Intelligence Scales for Children – 4 (Wechsler, 2003), as a measure of working memory; and the Tower of London subtest of the Delis-Kaplan Executive Functioning scales (Delis et al., 2001) as a measure of executive planning abilities. Autism diagnoses were confirmed in the autism group via administration of the Autism Diagnostic Observation Schedule, Module 3 (Lord et al., 2002) by trained graduate clinicians and the Social Communication Questionnaire (Rutter et al., 2003) parent report form, supplemented by clinical observation. Diagnostic evaluations were supervised by a licensed clinical psychologist (the first author) with extensive experience in autism evaluation. Non-autistic status was confirmed in the non-autistic comparison group using the Social Communication Questionnaire (Rutter et al., 2003). Adult participants ($n = 11$) completed a self-report questionnaire verifying an absence of developmental delays or concerns; they did not complete clinical assessments, which would have been age-inappropriate in most cases.

Speech-movement coordination.

A total of 3486 data points (beat gestures or speech peaks) were recorded. To minimize effects of outliers, we excluded data points that were equal to or greater than 3 SD from the mean (9.3% of the data). In some samples (8.4% of the data), there were no speech peaks detected between two gesture peaks, above the thresholded peak finding function (see below); these data were also excluded. The Amplitude Envelope was highly reliable, providing a viable method for tracking of acoustic peaks in speech. The remaining 2980 datapoints were included in all analyses.

Estimating movement peaks. Estimation of movement timing was partially automated. We calculated the instantaneous pixel change in video frames using a Python script provided by Brookshire and colleagues (Brookshire et al., 2017); from these values, we applied a first-order 10 Hz Butterworth filter for smoothing and then identified the derivatives of instantaneous pixel change (pixel change acceleration). Via visual inspection of the video data in ELAN (Lausberg & Sloetjes, 2009), we manually identified the time window encompassing the middle of each movement's extension phase to the middle of the subsequent flexion phase; thus, the time window included the point of maximum extension of the arm. We also checked for a peak in deceleration during that window, as registered by the pixel change acceleration time series viewable alongside the video data in ELAN (Crasborn et al., 2006). The exact timing of the peak in deceleration was extracted using a custom written R function using (R base) which, for each time window, assessed the moment of the maximum pixel change deceleration.

Identification of acoustic peaks in speech: Amplitude envelopes. Pilot analyses used the Fundamental Frequency estimation tool in PRAAT; however, with occasionally unreliable recording quality (e.g., when participants moved from the microphone or there were external noises), voicing was observed but F0 could not be tracked. Thus, peaks in the speech signal were identified via gross changes in the raw audio waveform (e.g., the Amplitude Envelope and its derivative, Amplitude Envelope Change), using a PRAAT script (He & Dellwo, 2016, 2017). We employed a Hilbert Transform to track gross amplitude changes in the speech signal, while ignoring finer structural changes; see Figure 2. The Amplitude Envelope is a scaled time series with values expressed in Hilbert units ranging from 0-1 (minimum to maximum observed amplitude). The Amplitude Envelope corresponds closely with the rhythmic structure of speech and with the oral kinematics of speech (Chandrasekaran et al., 2009; Tilsen & Arvaniti, 2013). In the current dataset, peaks in the amplitude envelope corresponded closely with each syllable of the rhyme, as shown in Figure 2; while this correspondence was generally true, where data were checked by hand, the relationship is not necessarily one-to-one (some syllables may have been omitted, and some peaks may not map onto a unique syllable). The peaks in the acceleration time series reflect downward and upward movements, each of which produced an upward positive acceleration and a deceleration (max flexion), and a downward positive acceleration and a deceleration (the downbeat). The current analysis included only the deceleration peaks produced during downbeats (which were initially hand-annotated).

To compute a temporal marker comparable to the gestural peaks, we calculated Amplitude Envelope Change as the time-derivative of the amplitude envelope, which represents changes in the speech amplitude envelope. Sudden positive changes ($SD > 33\%$) of the amplitude change, chosen by algorithm identification of peaks above the 33% threshold in the acoustic Amplitude Envelope Change time series, served as markers of speech segment onset Peaks, excluding smaller signal fluctuations. This threshold, though arbitrary, seemed to associate most closely with the syllable boundaries that it was aimed to detect, based on visual inspection of several audio samples.

Computing gesture-speech coordination and rhythmicity. Calculations of the temporal synchrony of motion peaks in the gesture signal relative to acoustic peaks in the speech signal utilized well-studied measures from the sensorimotor synchronization literature (Repp, 2005). Figure 2 provides a graphical overview of these measures, which evaluated *asynchrony* and speed and variation in *rhythmicity*.

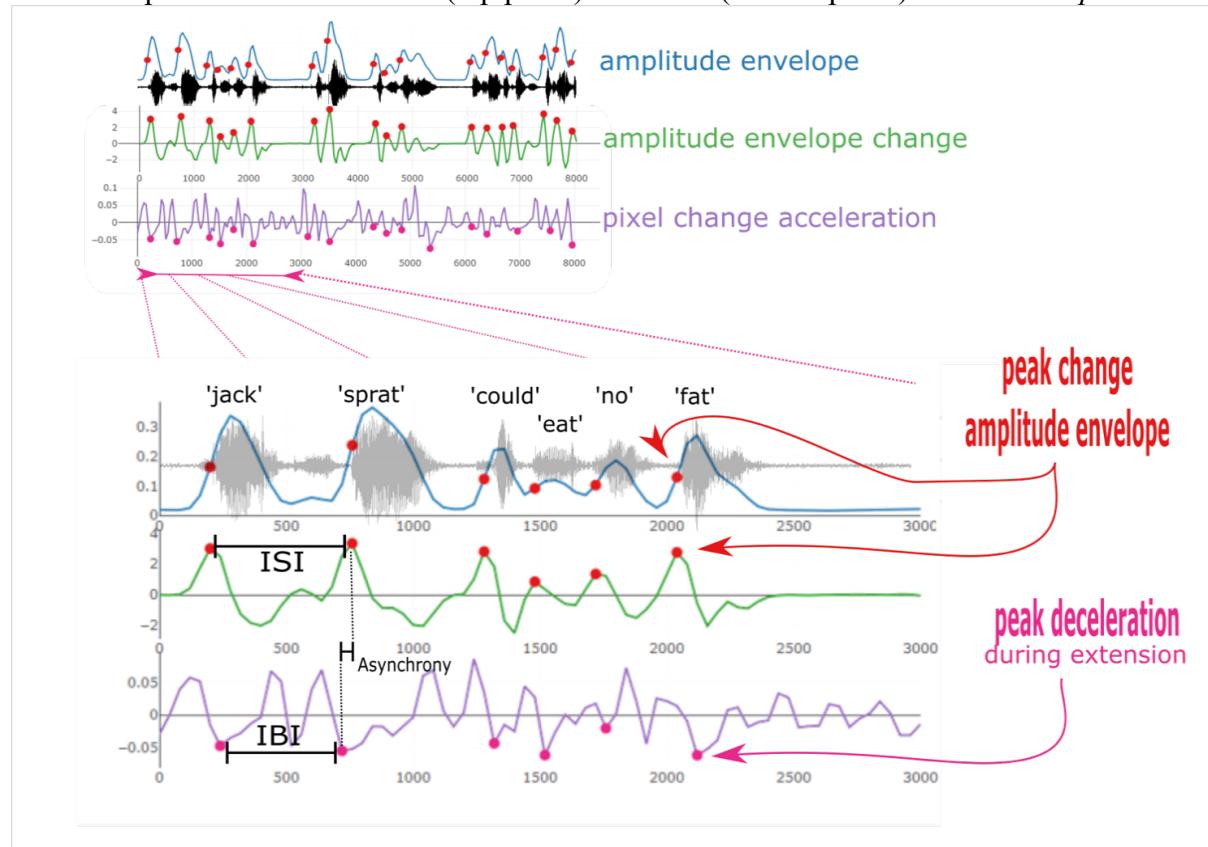
Asynchrony. Two measures of asynchrony were computed. For each movement peak deceleration, we identified the closest Amplitude Envelope Change peak. The difference in msec

between these peaks served as the estimation of temporal asynchrony of speech and movement. In addition, we assessed the stability (consistency) of speech-movement coordination, by calculating the variability across each participant's speech-movement peak asynchrony. Specifically, we calculated the standard deviation of peak asynchrony for each participant, and each trial; smaller SD values indicate greater stability of speech-movement coordination.

Inter-beat interval (IBI) and inter-speech interval (ISI). In addition to the coordination of individual gesture-speech events, the sensorimotor synchronization literature highlights the importance of the rhythmic qualities of speech-movement coordination (Michon et al., 2022; Repp, 2005). We measured the temporal distance between, or speed of, beat (IBI) or speech (ISI) segments, to calculate inter-beat interval (IBI) and inter-speech interval (ISI) measures. The standard deviation of these measures provided an estimate of the rigidity or rhythmicity of these intervals. For example, a lower standard deviation indicates that events were produced at more regular intervals.

Figure 2.
Graphical overview of gesture-speech coordination measures

Note. Sample time series of 8 sec (top panel) and 3 sec (bottom panel) of the *Jack Sprat* trial



produced by an adult participant. Blue lines (top row) = speech signal peaks. Green lines (middle row) = amplitude envelope changes. Purple lines (bottom row) = gestural motion peaks (pixel change acceleration). Y axes represent 0 to 1 scales (speech amplitude envelope) or Z-scores (gesture pixel change).

Results

Speech-movement asynchrony. Coordination was assessed using *nlme* mixed regression models (random intercept for participants) in R (Pinheiro et al., 2017; R Core Team, 2012). Basic models predicting overall means were contrasted with models including group (autism vs. neurotypical vs. adult) as a predictor; see Table 2 for descriptive statistics, and Figure 3 for a graphical overview. Results including significant effects of group were further investigated with Bonferroni corrected post-hoc *t*-tests implemented with the *lsmeans* package in R (Lenth & Lenth, 2018). Nursery rhyme was not included as a random factor due to data loss (see open science statement).

Participants across groups displayed negative mean asynchrony, shown in Figure 3; the apices (peaks) in hand movements reliably *preceded* speech peaks (that is, accented syllables). There were no reliable group differences for mean gesture-speech asynchrony; see Table 2, Figure 3. Flatter distributions indicate more variability around the mean. Groups differed in the variability of asynchrony (*SD*); speech and movements were less tightly coordinated in groups with higher variability (larger *SD*), shown as flatter distributions in Figure 3. Post-hoc comparisons indicated that only the adult/non-autistic adolescent difference was significant; the adult group had more consistent speech-movement coordination (an estimated difference of 26 ms, *SE* = 8.76, *p* = 0.02). The autistic adolescent group's asynchrony did not differ from the adults (an estimated difference of 17 ms, *SE* = 8.89, *p* = 0.18). Thus, there was more reliable gesture-speech coordination for adults as compared to neurotypical adolescents. This set of results suggests that the method and analysis were sufficiently sensitive to detect individual and group differences in coordination, even in this small sample, given the finding of a developmental change between the adolescent and adult neurotypical participants, with the adults showing tighter synchrony. Results indicate that the autistic group displayed speech-movement coordination equivalent to their age-matched non-autistic peers.

Table 2

Results of gesture-speech coordination analyses

	Autistic teens	Non-autistic teens	Adults	Mixed regression; change χ^2 [1]	Post-hoc <i>p</i> - corrected tests
Asynchrony	-22 (130) [-23, -41]	-32 (137) [-23, -41]	-37 (104) [-31, -43]	2.09, <i>p</i> = .351	No group differences
Speech-movement peak offset					
Asynchrony SD	121 (41) [109, 132]	129 (39) [119, 109]	102 (33) [94, 111]	8.21, <i>p</i> = .017	Adults less variable than non-autistic teens
Variability of speech-movement peak offset					
Inter-beat interval (IBI)	659 (325) [681, 638]	709 (370) [686, 733]	557 (279) [541, 573]	6.09, <i>p</i> = .048	No group differences
Tempo – movement peaks					
Inter-beat interval SD	277 (114) [315, 237]	277 (141) [240, 314]	229 (114) [200, 257]	2.39, <i>p</i> = .301	No group differences
Rhythmicity –					

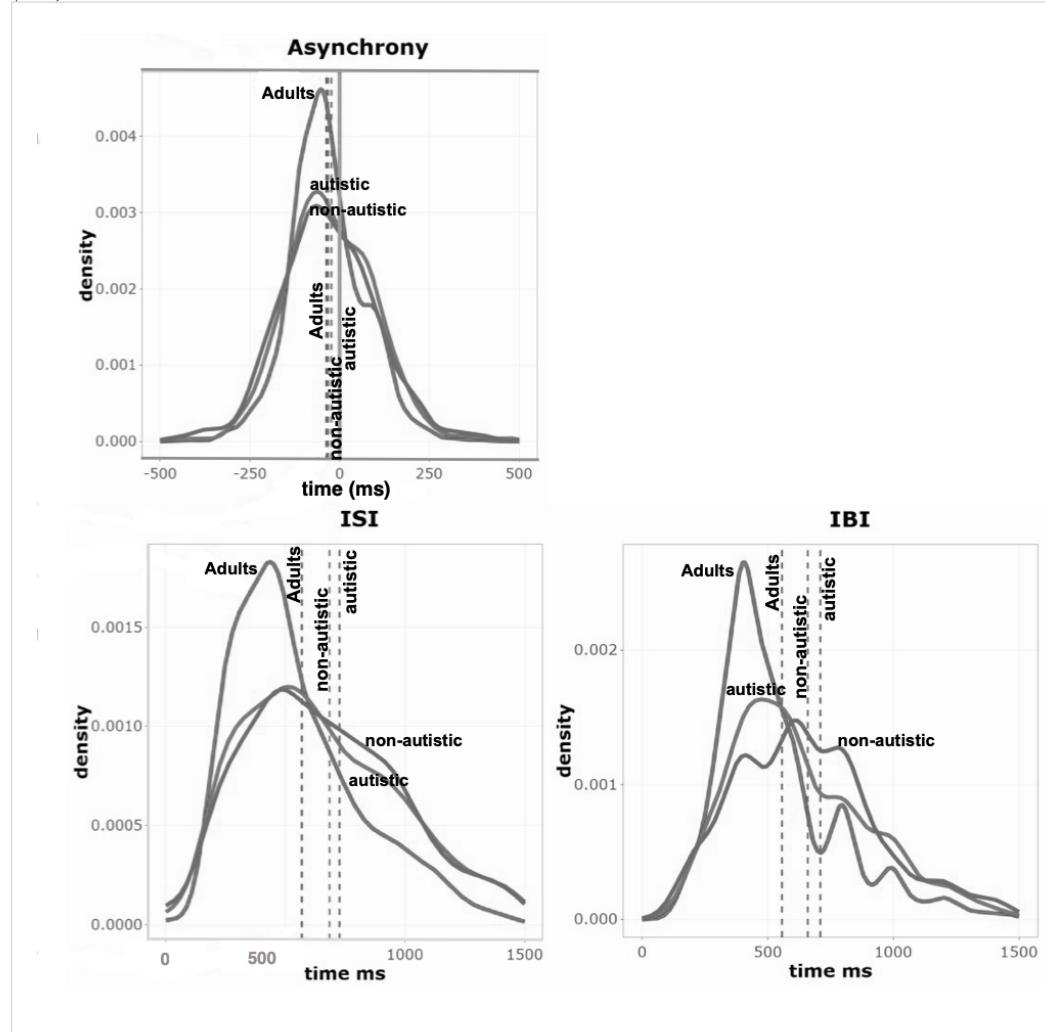
movement peaks					
Inter-speech interval (ISI)	685 (371) [660, 709]	725 (415) [699, 753]	568 (321) [550, 587]	7.73, $p = .021$	Adults faster than non-autistic teens
Tempo – speech peaks					
Inter-speech interval SD	341 (121) [308, 375]	352 (308) [322, 381]	290 (120) [261, 321]	7.74, $p = .020$	Adults steadier than non-autistic teens
Rhythmicity – speech peaks					

Note. Group results are presented as $M(SD)$, 95%CI [lower, upper]. All data are presented in ms.

Mixed regression model results represent change in χ^2 with Group as a predictor. Post-hoc comparisons are reported with Bonferroni corrections (significant at $p < .025$).

Figure 3.

Density distribution plots for asynchrony, inter-beat interval (IBI), and inter-syllable-interval (ISI).



Note. Dashed lines show group means. The y-axis indicates the smoothed density estimate representing the prevalence of an observed value on the x-axis. Asynchrony indicates the time in

milliseconds between the gesture peak and the speech speak, where negative values indicate that gesture is leading. ISI and IBI represent the time of intervals in ms of inter-speech and inter-gesture intervals, respectively.

Results of ISI and IBI analyses. The analysis of inter-syllable-intervals (ISI) yielded a significant effect of group; see Table 2. This indicates that groups differed in rhythmicity, or tempo, of speech peak production. Post-hoc comparisons indicated that the adults had a significantly shorter ISI than the non-autistic adolescents, indicating a faster tempo (estimated difference = 192 ms, $SE = 67.67$, $p = 0.025$). The autistic adolescents also displayed a slower tempo relative to adults, though the difference was not significant (estimated difference = 142 ms, $SE = 68.01$, $p = .138$). There were no reliable differences between autistic and non-autistic groups in ISI. Similarly, ISI variability (SD) showed a significant effect of group, indicating greater rhythmicity of speech peaks in adults relative to non-autistic adolescents (estimate difference = 26 ms, $SE = 8.76$, $p = .017$). Autistic adolescents had greater mean ISI SD compared to adults, but this difference was not significant (estimated difference = 18 ms, $SE = 8.89$, $p = .178$). There were no reliable differences between autistic and non-autistic groups on this measure. Thus, the adults maintained a much faster speech tempo, and a more rhythmic pace, as compared to non-autistic adolescents, and autistic adolescents' performance was between adults and non-autistic adolescents. There were no statistically significant effects of group in gesture intervals (IBI) or in IBI SD (rhythmicity). Exploratory correlational analyses examined the associations between working memory (digit span) and planning (Tower) abilities, on the one hand, and the synchrony and rhythmicity variables shown in Table 2. While results are necessarily unreliable given the sample size, they suggested no meaningful correlations (all p 's $> .10$).

These results suggest differences in speech-movement coordination for neurotypical adults as compared to younger adolescents, indicating developmental effects. However, there were no differences in coordination between autistic and non-autistic peers.

Discussion

Gesture-speech synchrony is temporally precise, with a strong coupling of gesture and speech prosody even in people with severe language production impairments (Jenkins & Pouw, 2023). A decrease in synchrony between spontaneous speech and iconic gestures has been documented in autism, with a negative impact on listener comprehension. That is, after watching a video of an individual telling a story, naive ratings of how easy to follow and clear the story was are strongly and significantly correlated with measures of speech-gesture synchrony (de Marchena & Eigsti, 2010). In the context of significant impairments in motor abilities in autism (Bhat, 2021), it is reasonable to propose that gesture-speech coordination differences are a function of impaired motor control; after all, both speech and gesture are motor actions and require ensemble coordination (Kelso et al., 1979). The current study was designed to test for difficulties in a task requiring the deliberate coordination of conscious, gesture-like hand movements, and speech. An absence of group differences on such a task would suggest that speech-gesture asynchrony reflects language production challenges, rather than motor difficulties (Kelly et al., 2010). In addition, we know little about developmental changes in gesture-speech coordination over the course of adolescence; we included adults in order to establish the mature endpoint of this developmental process.

To address these issues, the current exploratory study investigated the deliberate synchronization of speech with beat-like movements in a small group of autistic and non-autistic

adolescents and non-autistic adults. Participants were asked to move their dominant hand in a biphasic beat-like pattern, while reciting rhythmic utterances (nursery rhymes); the task was analogous to spontaneous co-speech gesture production, but with reduced *utterance planning* demands (because the utterances were read from a text prompt). Partially automated video and audio analysis tools were used to capture the primary peaks in both signals (that is, the endpoint of the beat-like hand movements, and the most prosodically prominent syllables in speech), and to measure the lag between these peaks. Analyses were thus designed to capture the coordination of two distinct motor acts: speech and gesture production.

Analyses of this conscious coordination task revealed that beat gestures were closely synchronized with speech across groups, and that hand movements reliably anticipated speech peaks. This negative mean asynchrony finding is consistent with results of other simple sensorimotor synchronization tasks such as finger tapping to a beep (Repp, 2005). Another interpretation is that, while the body movement precedes the speech peak, characteristics of the speech peak correlate with earlier-occurring characteristics of the movement, suggesting that the two are tightly bound together as a single communicative act; this was the conclusion suggested by a study of pitch accents and brow raises during spontaneous speech (Gast, 2023). This possibility calls for additional research.

Comparisons of autistic and non-autistic adolescents revealed no group differences in synchrony, tempo, or rhythmicity. This contrasts with results of prior studies reporting striking asynchrony of speech with spontaneously produced gestures with the autistic group showing significantly greater asynchrony (de Marchena & Eigsti, 2010; Morett et al., 2016). The difference between previous studies and the current one centers on language production; prior studies examined spontaneous narratives, whereas in the current study, participants read aloud a prepared text. Furthermore, the current study is not about semantic synchrony, but about kinematic-acoustic coupling, while the previous studies have looked at the synchrony between the lexical affiliate and the stroke (de Marchena & Eigsti, 2010; Morett et al., 2016). While preliminary, the current result is consistent with the hypothesis that linguistic factors may be more important contributors to asynchrony in autism, relative to motor coordination. That is, presented with a highly rhythmic phrases to read aloud, while performing a simple biphasic repeated hand motion, the autistic participants showed no difference in coordination.

Cognitive resource availability constrains language processing. For example, a study using a dual-task method with either low or high working memory demands reported differences in the perception of congruent or incongruent speech-gesture combinations (Momsen et al., 2020). The experimenter-defined (instructed) tasks in the current study presented minimal working memory and planning demands. If cognitive resource availability is the critical bottleneck, one would predict that asynchrony might vary as a function of morphosyntactic complexity and word frequency. This leads to a prediction that individuals with reduced cognitive resources (e.g., working memory) would show greater asynchrony, and could help explain the finding of speech-gesture asynchrony in autism, which is generally characterized by more limited verbal and visuospatial working memory (Ellis Weismer et al., 2018; Schuh & Eigsti, 2012; St John et al., 2022).

The results here are consistent with findings of spared rhythmic processing by autistic children in identifying beeps that were well aligned with the metric or rhythmic structure of music (Dahary et al., 2023) and broader findings of intact melodic and rhythmic perception in autistic children (Jamey et al., 2019). Together, these results, along with the current study, suggest that the highly rhythmic structure of nursery rhymes may have given the autistic

participants an advantage relative to spontaneous speech production. This suggestion awaits further study.

The current results indicate intact, or even developmentally advanced, synchrony for “prosodic,” non-representational, beat-like movements; in contrast, synchrony of representational (iconic) gestures seems to be disrupted in autism (De Marchena & Eigsti, 2010). Relevant to our understanding of the interface between spoken prosody and gesture production, the current study suggests the hypothesis that processing more complex conceptual semantic content, adds complexity to the coordination of gesture and speech.

A second finding in this study was the presence of differences between non-autistic adolescents and adults, indicating developmental increases in the precision of speech-movement coordination. Coordination in the autistic adolescents was intermediate between the other two groups, with no significant differences; the study is underpowered to detect subtle group differences. As one might expect, adults displayed more consistent timing between speech and movement (higher speech-movement coupling), a faster speech tempo, and greater rhythmicity in syllabic intervals, compared to non-autistic adolescents. On discrete timing indices, the autistic adolescent and adult groups did not differ, suggesting intact or even developmentally advanced explicit bimodal timing of self-produced actions. We did not evaluate puberty stages in this study; it is possible that participants in the autism group were physically advanced relative to their age-matched peers, as has been reported for some neurodevelopmental conditions (Siddiqi et al., 1999). The persistence of developmental effects in coordination in adolescence is somewhat surprising, given the early emergence of coordination of rhythmic arm movements and infant vocalizations (e.g., Ejiri & Masataka, 2001). Indeed, rudimentary vocal-motor synchronization is established in infancy (Ejiri & Masataka, 2001; Pouw & Fuchs, 2022). The finding that coordination continues to develop and improve during adolescence may highlight the additional cognitive constraints involved in coordinating motor action and utterance production.

Limitations. These results are limited in several dimensions. *First*, the labor-intensive nature of analyses limited inclusion to a small and relatively homogeneous sample of adolescent participants, spanning a relatively wide age range of 12-17 years; adult autistic participants were not included. Given the limited data at distinct ages, the results do not probe trajectories of development, and the results should be taken as exploratory; conclusive results await replication within a larger sample. *Second*, the autistic participants were not representative of the broader autism population, given their age-appropriate language and cognitive abilities; results might look very different with a more cognitively delayed sample. *Third*, the current sample was homogenous with respect to race and ethnicity, and composed of a larger number of males. *Finally*, executive processes might be engaged in an unexpected way by the artificial task demands; direct assessment of this possibility awaits further study.

The results of this preliminary study suggest that speech and movement coordination in a deliberate task is intact in autistic individuals, in contrast to findings of reduced coordination of spontaneous, self-generated representational gestures with speech during discourse. This finding is consistent with other reports that responses look more typical in autism when participants receive explicit guidance to perform a behavioral task (Eigsti, 2013). Results also suggest the importance of rhythmic action in the context of speech production, and point to the long arc of development in this domain, with differences between the adult state apparent even in neurotypical adolescents ages 12-17 years.

References

Alibali, M. W., Kita, S., & Young, A. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language & Cognitive Processes, 15*, 593-613.

Amaral, J. L., Romero, V., Kloos, H., & Richardson, M. J. (2017). Complexity and autism spectrum disorder: Exploring hysteresis in a grasping task. *AIMS Medical Science, 4*(1), 113-130.

Attwood, A., Frith, U., & Hermelin, B. (1988). The understanding and use of interpersonal gestures by autistic and Down's syndrome children. *Journal of Autism and Developmental Disorders, 18*(2), 241-257.

Bartak, L., Rutter, M., & Cox, A. (1975). A comparative study of infantile autism and specific developmental language disorder: I. The Children. *British Journal of Psychiatry, 126*, 127-145.

Bartak, L., Rutter, M., & Cox, A. (1977). A comparative study of infantile autism and specific developmental receptive language disorders III: Discriminant function analysis. *Journal of Autism and Childhood Schizophrenia, 7*(4), 383-396.

Bhat, A. N. (2021). Motor impairment increases in children with autism spectrum disorder as a function of social communication, cognitive and functional impairment, repetitive behavior severity, and comorbid diagnoses: A SPARK study report. *Autism Research, 14*(1), 202-219. <https://doi.org/10.1002/aur.2453>

Bhat, A. N., Landa, R. J., & Galloway, J. C. (2011). Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders. *Phys Ther, 91*(7), 1116-1129. <https://doi.org/10.2522/ptj.20100294>

Bloch, C., Tepest, R., Jording, M., Vogeley, K., & Falter-Wagner, C. M. (2022). Intrapersonal synchrony analysis reveals a weaker temporal coherence between gaze and gestures in adults with autism spectrum disorder. *Sci Rep, 12*(1), 20417. <https://doi.org/10.1038/s41598-022-24605-8>

Bono, M. A., Daley, L. T., & Sigman, M. (2004). Relations among joint attention, amount of intervention and language gain in autism. *Journal of Autism and Developmental Disorders, 34*(5), 495-505. <https://doi.org/10.1007/s10803-004-2545-x>

Brentari, D., & Goldin-Meadow, S. (2017). Language Emergence. *Annu Rev Linguist, 3*, 363-388. <https://doi.org/10.1146/annurev-linguistics-011415-040743>

Brookshire, G., Lu, J., Nusbaum, H. C., Goldin-Meadow, S., & Casasanto, D. (2017). Visual cortex entrains to sign language. *Proceedings of the National Academy of Sciences, 114*(24), 6352-6357. <https://doi.org/10.1073/pnas.1620350114>

Butera, C., Delafield-Butt, J., Kilroy, E., Harrison, L., Anzulewicz, A., Sobota, K., & Aziz-Zadeh, L. (2017, 2017-04-05). *Potential movement biomarkers for autism in children and adolescents* University of Southern California Research Day 2017, University of Southern California.

Camaioni, L., Peruccini, P., Muratori, F., & Milone, A. (1997). Brief report: a longitudinal examination of the communicative gestures deficit in young children with autism. *Journal of Autism and Developmental Disorders, 27*(6), 715-725.

Cannon, J., Eldracher, E., Cardinaux, A., Irfan, F., Bungert, L., Li, C., O'Brien, A., Treves, I., Diamond, S., & Sinha, P. (2023). Rhythmic and interval-based temporal orienting in autism. *Autism Res, 16*(4), 772-782. <https://doi.org/10.1002/aur.2892>

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., & Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proc Natl Acad Sci U S A*, 104(45), 17825-17830. <https://doi.org/0706273104> [pii] 10.1073/pnas.0706273104 [doi]

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Comput Biol*, 5(7), e1000436. <https://doi.org/10.1371/journal.pcbi.1000436>

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, 30(1), 213-236.

Colgan, S. E., Lanter, E., McComish, C., Watson, L. R., Crais, E. R., & Baranek, G. T. (2006). Analysis of social interaction gestures in infants with autism. *Child Neuropsychology*, 12, 307-319.

Cook, J. L., Blakemore, S. J., & Press, C. (2013). Atypical basic movement kinematics in autism spectrum conditions. *Brain*, 136(Pt 9), 2816-2824. <https://doi.org/10.1093/brain/awt208>

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, 106(2), 1047-1058. [https://doi.org/S0010-0277\(07\)00114-X](https://doi.org/S0010-0277(07)00114-X) [pii]

10.1016/j.cognition.2007.04.010

Crasborn, O., Sloetjes, H., Auer, E., & Wittenburg, P. (2006). Combining video and numeric data in the analysis of sign languages with the ELAN annotation software. Proceedings of the 2nd Workshop on the Representation and Processing of Sign languages: Lexicographic matters and didactic scenarios,

Dahary, H., Rimmer, C., & Quintin, E.-M. (2023). Musical Beat Perception Skills of Autistic and Neurotypical Children. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-022-05864-w>

de Marchena, A., & Eigsti, I. M. (2010). Conversational gestures in autism spectrum disorders: Asynchrony but not decreased frequency. *Autism Research*, 3, 311-322. <https://doi.org/10.1002/aur.159>

de Marchena, A., & Eigsti, I. M. (2014). Context counts: The impact of social context on gesture rate in verbally fluent adolescents with autism spectrum disorder. *Gesture*, 14(3), 375–393. <https://doi.org/10.1075/gest.14.3.05mar>

de Marchena, A., Kim, E. S., Bagdasarov, A., Parish-Morris, J., Maddox, B. B., Brodkin, E. S., & Schultz, R. T. (2019). Atypicalities of Gesture Form and Function in Autistic Adults. *J Autism Dev Disord*, 49(4), 1438-1454. <https://doi.org/10.1007/s10803-018-3829-x>

Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan Executive Function System*. Psychological Corporation.

Dimitrova, N., & Özçalışkan, Ş. (2022). Identifying Patterns of Similarities and Differences between Gesture Production and Comprehension in Autism and Typical Development. *J Nonverbal Behav*, 46(2), 173-196. <https://doi.org/10.1007/s10919-021-00394-y>

Dziuk, M. A., Gidley Larson, J. C., Apostu, A., Mahone, E. M., Denckla, M. B., & Mostofsky, S. H. (2007). Dyspraxia in autism: association with motor, social, and communicative deficits. *Developmental Medicine and Child Neurology*, 49(10), 734-739. <https://doi.org/DMCN734> [pii] 10.1111/j.1469-8749.2007.00734.x

Eigsti, I. M. (2013). A review of embodiment in autism spectrum disorders. *Frontiers in psychology*, 4(224), 10.3389/fpsyg.2013.00224. eCollection 02013.

Eigsti, I. M. (2024). The Autism Constellation and Neurodiversity: Long-Term and Adult Outcomes in Autism Spectrum Disorder. *Pediatr Clin North Am*, 71(2), 327-341. <https://doi.org/10.1016/j.pcl.2024.01.003>

Ejiri, K., & Masataka, N. (2001). Co-occurrence of preverbal vocal behavior and motor action in early infancy. *Developmental Science*, 4(1), 40–48. <https://doi.org/10.1111/1467-7687.00147>

Ellis Weismier, S., Kaushanskaya, M., Larson, C., Mathée, J., & Bolt, D. (2018). Executive Function Skills in School-Age Children With Autism Spectrum Disorder: Association With Language Abilities. *J Speech Lang Hear Res*, 61(11), 2641-2658. https://doi.org/10.1044/2018_jslhr-1-rsaut-18-0026

Fabbri-Destro, M., Cattaneo, L., Boria, S., & Rizzolatti, G. (2009). Planning actions in autism. *Experiments in Brain Research*, 192(3), 521-525. <https://doi.org/10.1007/s00221-008-1578-3> [doi]

Fuhrmann, D., Knoll, L. J., & Blakemore, S.-J. (2015). Adolescence as a Sensitive Period of Brain Development. *Trends in Cognitive Sciences*, 19(10), 558-566. <https://doi.org/10.1016/j.tics.2015.07.008>

Garcia-Perez, R. M., Lee, A., & Hobson, R. P. (2007). On intersubjective engagement in autism: A controlled study of nonverbal aspects of conversation. *Journal of Autism and Developmental Disorders*, 37(7), 1310-1322. Epub 2006 Nov 1314.

Gast, V. (2023). The Temporal Alignment of Speech-Accompanying Eyebrow Movement and Voice Pitch: A Study Based on Late Night Show Interviews. *Behav Sci (Basel)*, 13(1). <https://doi.org/10.3390/bs13010052>

Gidley Larson, J. C., Bastian, A. J., Donchin, O., Shadmehr, R., & Mostofsky, S. H. (2008). Acquisition of internal models of motor tasks in children with autism. *Brain*, 131(Pt 11), 2894-2903. <https://doi.org/awn226> [pii] 10.1093/brain/awn226 [doi]

Glazebrook, C., Gonzalez, D., Hansen, S., & Elliott, D. (2009). The role of vision for online control of manual aiming movements in persons with autism spectrum disorders. *Autism*, 13(4), 411-433. <https://doi.org/13/4/411> [pii] 10.1177/1362361309105659 [doi]

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20(3), 267-272. <https://doi.org/PSCI2297> [pii]

10.1111/j.1467-9280.2009.02297.x

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends Cogn Sci*, 9(5), 234-241. <https://doi.org/10.1016/j.tics.2005.03.006>

Habets, B., Kita, S., Shao, Z., Ozyurek, A., & Hagoort, P. (2010). The role of synchrony and ambiguity in speech-gesture integration during comprehension. *Journal of Cognitive Neuroscience*. <https://doi.org/10.1162/jocn.2010.21462> [doi]

Hamilton, A. F., Brindley, R. M., & Frith, U. (2007). Imitation and action understanding in autistic spectrum disorders: how valid is the hypothesis of a deficit in the mirror neuron system? *Neuropsychologia*, 45(8), 1859-1868. [https://doi.org/S0028-3932\(06\)00478-7](https://doi.org/S0028-3932(06)00478-7) [pii] 10.1016/j.neuropsychologia.2006.11.022

Harrison, L. A., Kats, A., Kilroy, E., Butera, C., Jayashankar, A., Keles, U., & Aziz-Zadeh, L. (2021). Motor and sensory features successfully decode autism spectrum disorder and combine with the original RDoC framework to boost diagnostic classification. *Sci Rep*, 11(1), 7839. <https://doi.org/10.1038/s41598-021-87455-w>

He, L., & Dellwo, V. (2016). A PRAAT-based algorithm to extract the amplitude envelope and temporal fine structure using the Hilbert transform. *Proceedings Interspeech*,

He, L., & Dellwo, V. (2017). Amplitude envelope kinematics of speech signal: parameter extraction and applications. In J. Trouvain, I. Steiner, & B. Möbius (Eds.), *Elektronische Sprachsignalverarbeitung* (pp. 1-8). TUDpress.

Hughes, C. (1996). Brief report: Planning problems in autism at the level of motor control. *Journal of Autism and Developmental Disorders*, 26, 99-107.

Ingersoll, B., & Lalonde, K. (2010). The impact of object and gesture imitation training on language use in children with autism spectrum disorder. *J Speech Lang Hear Res*, 53(4), 1040-1051. [https://doi.org/10.1044/1092-4388\(2009/09-0043\)](https://doi.org/10.1044/1092-4388(2009/09-0043))

Jamey, K., Foster, N. E. V., Sharda, M., Tuerk, C., Nadig, A., & Hyde, K. L. (2019). Evidence for intact melodic and rhythmic perception in children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, 64, 1-12. <https://doi.org/https://doi.org/10.1016/j.rasd.2018.11.013>

Jansiewicz, E. M., Goldberg, M. C., Newschaffer, C. J., Denckla, M. B., Landa, R., & Mostofsky, S. H. (2006). Motor signs distinguish children with high functioning autism and Asperger's syndrome from controls. *Journal of Autism and Developmental Disorders*, 36(5), 613-621. <https://doi.org/10.1007/s10803-006-0109-y> [doi]

Jenkins, T., & Pouw, W. (2023). Gesture-speech coupling in persons with aphasia: A kinematic-acoustic analysis. *J Exp Psychol Gen*, 152(5), 1469-1483. <https://doi.org/10.1037/xge0001346>

Kadavá, Š., Cwiek, A., Fuchs, S., & Pouw, W. (August 7-8, 2023, 7-8 August 2023). Is gesture-speech physics at work in rhythmic pointing? Evidence from Polish counting-out rhymes. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)*, Prague.

Keating, C. T., Hickman, L., Leung, J., Monk, R., Montgomery, A., Heath, H., & Sowden, S. (2023). Autism-related language preferences of English-speaking individuals across the globe: A mixed methods investigation. *Autism Res*, 16(2), 406-428. <https://doi.org/10.1002/aur.2864>

Kelly, S. D., Ozyurek, A., & Maris, E. (2010). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychological Science*, 21(2), 260-267. <https://doi.org/0956797609357327> [pii], 10.1177/0956797609357327 [doi]

Kelso, J. A., Southard, D. L., & Goodman, D. (1979). On the nature of human interlimb coordination. *Science*, 203(4384), 1029-1031. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=424729

Kemper, G. L., van der Sluis, A., Brink, M. S., Visscher, C., Frencken, W. G., & Elferink-Gemser, M. T. (2015). Anthropometric Injury Risk Factors in Elite-standard Youth Soccer. *Int J Sports Med*, 36(13), 1112-1117. <https://doi.org/10.1055/s-0035-1555778>

Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57(3), 396-414.

Lau, J. C. Y., Losh, M., & Speights, M. (2023). Differences in speech articulatory timing and associations with pragmatic language ability in autism. *Res Autism Spectr Disord*, 102. <https://doi.org/10.1016/j.rasd.2023.102118>

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers*, 41(3), 841-849. <https://doi.org/10.3758/BRM.41.3.841>

Lenth, R., & Lenth, M. R. (2018). Package 'lsmeans'. *The American Statistician*, 34(4), 216-221.

Levelt, W. J. M., Richardson, G., & la Heij, W. (1985). Pointing and voicing in deictic expressions. *Journal of Memory and Language*, 24, 133-164.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2002). *Autism Diagnostic Observation Schedule (ADOS)*. Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule: ADOS-2*. Western Psychological Services Los Angeles, CA.

Lord, C., Rutter, M., & LeCouteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659-685.

Loveland, K. A., & Landry, S. H. (1986a). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders*, 16(3), 335-349. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citati on&list_uids=3558291

Loveland, K. A., & Landry, S. H. (1986b). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders*, 16, 335-349.

Luyster, R., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of Child Language*, 34(3), 623-654.

Macneil, L. K., & Mostofsky, S. H. (2012). Specificity of dyspraxia in children with autism. *Neuropsychology*. <https://doi.org/10.1037/a0026955> [pii] 10.1037/a0026955 [doi]

McAuliffe, D., Pillai, A. S., Tiedemann, A., Mostofsky, S. H., & Ewen, J. B. (2017). Dyspraxia in ASD: Impaired coordination of movement elements. *Autism Res*, 10(4), 648-652. <https://doi.org/10.1002/aur.1693>

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92, 350-371.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Michon, M., Zamorano-Abramson, J., & Aboitiz, F. (2022). Faces and Voices Processing in Human and Primate Brains: Rhythmic and Multimodal Mechanisms Underlying the Evolution and Development of Speech. *Frontiers in psychology*, 13, 829083. <https://doi.org/10.3389/fpsyg.2022.829083>

Momsen, J., Gordon, J., Wu, Y. C., & Coulson, S. (2020). Verbal working memory and co-speech gesture processing. *Brain Cogn*, 146, 105640. <https://doi.org/10.1016/j.bandc.2020.105640>

Morett, L. M., O'Hearn, K., Luna, B., & Ghuman, A. S. (2016). Altered gesture and speech production in ASD detract from in-person communicative quality. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-015-2645-9>

Mostofsky, S. H., & Ewen, J. B. (2011). Altered connectivity and action model formation in autism Is autism. *Neuroscientist*, 5, 5.

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.

Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: The contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry*, 27, 657-669.

Murillo, E., Camacho, L., & Montero, I. (2021). Multimodal Communication in Children with Autism Spectrum Disorder and Different Linguistic Development. *J Autism Dev Disord*, 51(5), 1528-1539. <https://doi.org/10.1007/s10803-020-04637-7>

Nebel, M. B., Eloyan, A., Nettles, C. A., Sweeney, K. L., Ament, K., Ward, R. E., Choe, A. S., Barber, A. D., Pekar, J. J., & Mostofsky, S. H. (2016). Intrinsic Visual-Motor Synchrony Correlates With Social Deficits in Autism. *Biol Psychiatry*, 79(8), 633-641. <https://doi.org/10.1016/j.biopsych.2015.08.029>

Nobe, S. (2010). Where do most spontaneous representational gestures actually occur with respect to speech? In D. McNeill (Ed.), *Language and Gesture* (pp. 186–198). Cambridge University Press. <https://doi.org/DOI:10.1017/CBO9780511620850.012>

Özçalışkan, Ş., Adamson, L. B., Dimitrova, N., & Baumann, S. (2017). Early gesture provides a helping hand to spoken vocabulary development for children with autism, down syndrome, and typical development. *Journal of Cognition and Development*, 18, 325–337. <https://doi.org/10.1080/15248372.2017.1329735>

Parrell, B., Goldstein, L., Lee, S., & Byrd, D. (2014). Spatiotemporal coupling between speech and manual motor actions. *Journal of Phonetics*, 42, 1-11.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). Package ‘nlme’. *Linear and nonlinear mixed effects models*, 3(1), 336.

Pouw, W., & Dixon, J. A. (2019). Entrainment and modulation of gesture-speech synchrony under delayed auditory feedback. *Cognitive science*, 43(3), e12721. <https://doi.org/10.1111/cogs.12721>

Pouw, W., & Fuchs, S. (2022). Origins of vocal-entangled gesture. *Neurosci Biobehav Rev*, 141, 104836. <https://doi.org/10.1016/j.neubiorev.2022.104836>

Pouw, W., Harrison, S. J., & Dixon, J. A. (2022). The importance of visual control and biomechanics in the regulation of gesture-speech synchrony for an individual deprived of proprioceptive feedback of body position. *Sci Rep*, 12(1), 14775. <https://doi.org/10.1038/s41598-022-18300-x>

Pouw, W., Trujillo, J. P., & Dixon, J. A. (2020). The quantification of gesture-speech synchrony: A tutorial and validation of multimodal data acquisition using device-based and video-based motion tracking. *Behav Res Methods*, 52(2), 723-740. <https://doi.org/10.3758/s13428-019-01271-9>

Purpura, G., Costanzo, V., Chericoni, N., Puopolo, M., Scattoni, M. L., Muratori, F., & Apicella, F. (2017). Bilateral Patterns of Repetitive Movements in 6- to 12-Month-Old Infants with Autism Spectrum Disorders. *Frontiers in psychology*, 8, 1168. <https://doi.org/10.3389/fpsyg.2017.01168>

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012. In: ISBN 3-900051-07-0.

Ramos-Cabo, S., Vulchanov, V., & Vulchanova, M. (2019). Gesture and Language Trajectories in Early Development: An Overview From the Autism Spectrum Disorder Perspective. *Frontiers in psychology*, 10, 1211. <https://doi.org/10.3389/fpsyg.2019.01211>

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychon Bull Rev*, 12(6), 969-992. <https://www.ncbi.nlm.nih.gov/pubmed/16615317>

Rimmer, C., Dahary, H., & Quintin, E. M. (2023). Links between musical beat perception and phonological skills for autistic children. *Child Neuropsychol*, 1-20. <https://doi.org/10.1080/09297049.2023.2202902>

Rinehart, N. J., Bradshaw, J. L., Brereton, A. V., & Tonge, B. J. (2001). Movement preparation in high-functioning autism and Asperger disorder: a serial choice reaction time task involving motor reprogramming. *Journal of Autism and Developmental Disorders*, 31(1), 79-88.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 31, 131-144.

Rohrer, P. L., Delais-Roussarie, E., & Prieto, P. (2023). Visualizing prosodic structure: Manual gestures as highlighters of prosodic heads and edges in English academic discourses. *Lingua*, 293, 103583.

Romero, V., Fitzpatrick, P., Roulier, S., Duncan, A., Richardson, M. J., & Schmidt, R. C. (2018). Evidence of embodied social competence during conversation in high functioning children with autism spectrum disorder. *PLoS One*, 13(3), e0193906. <https://doi.org/10.1371/journal.pone.0193906>

Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire (SCQ)*. Western Psychological Services.

Scharoun, S. M., & Bryden, P. J. (2016). Anticipatory Planning in Children with Autism Spectrum Disorder: An Assessment of Independent and Joint Action Tasks. *Front Integr Neurosci*, 10, 29. <https://doi.org/10.3389/fnint.2016.00029>

Schuh, J. M., & Eigsti, I. M. (2012). Working memory, language skills, and autism symptomatology. *Behavioral Sciences*, 2, 207-218. <https://doi.org/10.3390/bs2040207>

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals* (4th ed.). Harcourt Assessment, Inc.

Siddiqi, S. U., Van Dyke, D. C., Donohoue, P., & McBrien, D. M. (1999). Premature sexual development in individuals with neurodevelopmental disabilities. *Developmental Medicine and Child Neurology*, 41(6), 392-395.

Silverman, L. B., Eigsti, I. M., & Bennetto, L. (2017). I tawt I taw a puddy tat: Gestures in Canary Row narrations by high-functioning adolescents with autism spectrum disorder. *Autism Research*, 10(8), 1353-1363. <https://doi.org/10.1002/aur.178>

So, W. C., Lui, M., Wong, T. K., & Sit, L. T. (2015). The use of hand gestures to communicate about nonpresent objects in mind among children with autism spectrum disorder. *J Speech Lang Hear Res*, 58(2), 373-382. https://doi.org/10.1044/2015_jslhr-1-14-0213

St John, T., Woods, S., Bode, T., Ritter, C., & Estes, A. (2022). A review of executive functioning challenges and strengths in autistic adults. *Clin Neuropsychol*, 36(5), 1116-1147. <https://doi.org/10.1080/13854046.2021.1971767>

Stevens, M. C., Fein, D. A., Dunn, M., Allen, D., Waterhouse, L. H., Feinstein, C., & Rapin, I. (2000). Subgroups of children with autism by cluster analysis: a longitudinal examination. *Journal of the American Academic of Child and Adolescent Psychiatry*, 39(3), 346-352. <https://doi.org/10.1097/00004583-200003000-00017>

Sutera, S., Pandey, J., Esser, E. L., Rosenthal, M. A., Wilson, L. B., Barton, M., Green, J., Hodgson, S., Robins, D. L., Dumont-Mathieu, T., & Fein, D. (2007). Predictors of optimal outcome in toddlers diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37(1), 98-107. <https://doi.org/10.1007/s10803-006-0340-6> [doi]

Tager-Flusberg, H., Calkins, S., Nolin, T., Baumberger, T., Anderson, M., & Chadwick-Dias, A. (1990). A longitudinal study of language acquisition in autistic and Down syndrome children. *Journal of Autism and Developmental Disorders*, 20(1), 1-21.

Tantam, D., Holmes, D., & Cordess, C. (1993). Nonverbal expression in autism of Asperger type. *Journal of Autism and Developmental Disorders*, 23(1), 111-133.

Taverna, E. C., Huedo-Medina, T. B., Fein, D. A., & Eigsti, I. M. (2021). The interaction of fine motor, gesture, and structural language skills: The case of autism spectrum disorder. *Res Autism Spectr Disord*, 86. <https://doi.org/10.1016/j.rasd.2021.101824>

Tilson, S., & Arvaniti, A. (2013). Speech rhythm analysis with decomposition of the amplitude envelope: Characterizing rhythmic patterns within and across languages. *The Journal of the Acoustical Society of America*, 134(1), 628-639. <https://doi.org/10.1121/1.4807565>

Trujillo, J. P., Özyürek, A., Kan, C. C., Sheftel-Simanova, I., & Bekkering, H. (2021). Differences in the production and perception of communicative kinematics in autism. *Autism Res*, 14(12), 2640-2653. <https://doi.org/10.1002/aur.2611>

van Swieten, L. M., van Bergen, E., Williams, J. H., Wilson, A. D., Plumb, M. S., Kent, S. W., & Mon-Williams, M. A. (2010). A test of motor (not executive) planning in developmental coordination disorder and autism. *Journal of Experimental Psychology: Human Perception and Performance*, 36(2), 493-498.

Vilà-Giménez, I., & Prieto, P. (2020). Encouraging kids to beat: Children's beat gesture production boosts their narrative performance. *Dev Sci*, 23(6), e12967. <https://doi.org/10.1111/desc.12967>

Vilensky, J. A., Damasio, A. R., & Maurer, R. G. (1981). Gait disturbances in patients with autistic behavior: a preliminary study. *Archives of Neurology*, 38(10), 646-649.

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209-232.

Wang, L., & Chu, M. (2013). The role of beat gesture and pitch accent in semantic processing: an ERP study. *Neuropsychologia*, 51, 2847-2855.

Wechsler, D. (2003). *Wechsler Intelligence Scales for Children- 4th Edition*. Psychological Corporation.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence* (Second ed.). Pearson Assessments.

Wetherby, A. M., & Prutting, C. A. (1984). Profiles of communicative and cognitive-social abilities in autistic children. *Journal of Speech and Hearing Research*, 27, 364-377.

Wong, M. K., & So, W. C. (2018). Absence of delay in spontaneous use of gestures in spoken narratives among children with Autism Spectrum Disorders. *Res Dev Disabil*, 72, 128-139. <https://doi.org/10.1016/j.ridd.2017.11.004>