Education, Language and Sociology Research
ISSN 2690-3644 (Print) ISSN 2690-3652 (Online)
Vol. 7, No. 1, 2026
www.scholink.org/ojs/index.php/elsr

Original Paper

Multimodal Parent-child Interactions in Autism: What Can We Learn from Them?

Yang Wang[1]

[1] Cognitive Science and Allied Health School, Beijing Language and Culture University, Beijing, China

Received: February 17, 2026 Accepted: March 2, 2026 Online Published: March 19, 2026

doi:10.22158/elsr.v7n1p119 URL: http://dx.doi.org/10.22158/elsr.v7n1p119

Abstract

Purpose: *Exploring the multimodal communicative features in autistic parent-child interactions.* ***Methods:*** *Based on Sigrid Norris'multimodal interaction theory, editing and annotating parent-child interactions in five autism-related documentaries, this study analyzed the modality configurations and structural relationships of individuals with autism in parent-child interactions.* ***Results:*** *It was found that children with autism use more nonverbal than verbal modalities in parent-child interactions and they are capable of selecting distinct modalities and modal structures for communication purposes. In addition, autistic children are prone to articulate their thoughts, attitudes, and emotions in parent-child interactions.* ***Conclusion:*** *These data reveal that we can foster deeper cognitive skills of autistic children, such as identity and emotional awareness in parent-child interactions and consider the communicative behaviors of autistic people through a differential rather than a defective perspective.*

Keywords

Autism, parents-children interaction, multi-modal, discourse analysis

1. Introduction

Generally, impaired social interaction and repetitive stereotypic behaviors are the main features of ASD (autism spectrum disorders), which are evident in early parent-child interactions in ASD children (Ingersoll, 2016). Meek et al. (2012) found that children with autism exhibited behavioral difficulties in early parent-child interactions in a variety of areas such as attention, interest expression, cooperation, and positive affect, and the development of these abilities is extremely essential to children's social interaction. In addition, parent-child intervention therapy, as one of the intensive behavioral interventions, not only alleviates the imbalance between the supply and demand of professional therapists and reduces the cost of interventions, but also enables the parent-child relationship to be fully utilized and interventions to be carried out as early as possible. Consequently, a comprehensive understanding of the characteristics in

parent-child interactions and an effective intervention approach based on them are crucial for the rehabilitation and development of autism.

1.1 Parent-child interaction in autism

Parent-child interaction refers to the mutual interaction activities between parents and children. Current research on parent-child interaction in children with ASD mainly addresses the effects and assessment, the significance, the influences, and the characteristics of parent- child interaction. In an exploration of the significance of parent-child interaction, early parent-child interaction in individuals with ASD is actually closely related to their later development (El-Ghoroury et al., 1999) and parent-child interaction not only positively influences the recovery of children with autism in the short term, but also predicts to some extent the future social interaction skills of autism (Wan et al., 2019, Siller & Sigman, 2002). It has been shown that parental emotional support and parent-child cohesion positively influence children's social skills (Haven et al., 2014) and that parent-child interaction therapy (PCIT), designed based on parent-child interaction, is effective in addressing behavioral difficulties in autism (Scudder et al., 2019; Zlomke et al., 2020).

Parent-child interaction is of great value in the treatment and rehabilitation of autism, but its characteristics have not been explored much. Studies on the characteristics of parent- child interactions in autism revealed differences in the total amount of gesture use, and the type of gesture use during spontaneous mother-child interactions in children with ASD, children with Down syndrome, and typically developing children, and based on this, it was proposed that the study of gestures in children with autism spectrum disorders could potentially assist us in identifying different phenotypes of autism, which may potentially result in the development of new therapies (Mastrogiuseppe et al., 2015). In terms of verbal comprehension, based on the performance of children with high-functioning autism (HFA) and children with asperger's syndrome (AS) in response to everyday questions posed by family members during dinner interactions, Kreme(2004) identified the opposite conclusion to previous cognitive psychology studies, namely that most human children were able to perceive the communicative intentions of their interlocutors and respond in a relevant way, which could be attributed to the different pragmatic strategies used by the parents.

On the basis of the above related literature, we can identify that parent-child interaction in autism is instrumental in the emotional development, language enhancement, psychological and cognitive development, and social interaction of children with ASD, however, current research has not conducted extensive in-depth studies on the characteristics of it, which lacks a clearer mechanism for parent-child intervention therapy and is not conducive to quantifying the intervention modalities and developing more effective intervention methods.

1.2 Multimodal Interaction Characteristics in Autism

The expression and transmission of meaning in social interactions is multimodal and demands a diversity of symbolic resources, as the case for parent-child interactions in autism. Apart from verbal behaviors, nonverbal behaviors, such as gaze, facial expressions, gestures, body posture, and scene layout, can also

be highly conducive for communication by conveying the speaker's intention or interest as well as clarifying or emphasizing thoughts (Landa, 2007).

It has been found that individuals with ASD exhibit various atypicalities in language. For example, they tend to have difficulty in initiating interactions and making conversational offers (Loveland & Tunali, 2005); ans frequently produce unintentionally offensive discourse and are unable to distinguish between old and new information (Baltaxe 1977). In addition, their understanding of others' speech may also be limited, especially in understanding metaphors or ambiguous language, such as humor or sarcasm (Happé, 1993). It has become clear that this atypicality affects the quality of conversations between individuals with autism and others (Fine et al., 1994; Loveland & Tunali, 2005).

It is widely assumed that early gesture use is associated with later language ability and is considered a critical precursor to early language production in normal healthy children (Bates & Dick, 2002; Iverson & Goldin, 2005). In terms of frequency and type of gesture use, ASD infants and children have a low frequency of gesture use and a delay in gesture recognition and comprehension (Mitchell et al., 2006) (Baranek, 1999; Ham et al., 2010; Smith & Bryson, 2007). Compared to their normally developing peers, young children with ASD produce fewer instructional gestures and fewer synchronous gestures (De Marchena & Eigsti, 2010), whereas the use of object gestures is similar to that of normal children. Contrary to ASD infants and toddlers, ASD children aged 4-13 years rely more on gestures in communication and comply more readily with nonverbal than verbal requests in comparison to normally developing children (Kurt, 2011; Smith & Bryson, 2007).

In regard to facial expressions, there was evidence that the facial expressions and vocal rhythms of individuals with autism are less natural than those of normal individuals (Grossman et al., 2013). In terms of postural modalities, Huang (2018) noted that autistic children suffer from some deficits in body perception and understanding and inability to recognize the emotions embodied in body postures. Chen and Chen (2022) explored head movement deficits in children with autism and found that head rotation movements (translational and rotational movements) of autistic children in social interactions were synchronized with their peers, but at lower levels of synchronization than typically developing children, which likely implied that they showed less attention to their interacting partners.

In summary, researches on autistic interactions has concentrated more on single modalities, such as verbal modality or gestural modality, and fewer comprehensive studies on multimodalities. In terms of research perspective, the majority of studies were conducted on an experimental paradigm in psychology rather than a discourse analysis paradigm grounded in interactive linguistics and social constructivism, which renders it challenging to demonstrate the overall picture of children with autism in real social interactions.

1.3 Multimodal Interaction Analysis Theory

In recent years, various theories of multimodal discourse analysis have gradually emerged as a hot topic for academia. Among them, the multimodal interaction analysis method proposed by Sigrid Norris is considered as a holistic analysis of multiple real-time, continuous and simultaneous communication processes of participants. With the mediated language theory and practice node analysis as the basis, and

activities as the basic analysis unit, it combines interactive sociolinguistics, sociosemiotics and multimodal discourse analysis to achieve a dynamic analysis of the interaction process (Zhang, 2016). In multimodal interaction analysis, each resource signs used in the interaction process carries a specific meaning and either modality has the potential to have a dominant position.

This argument has been confirmed in various fields such as social interaction, teaching activities, and politics (Norris, 2007; Ticca, 2017; Steger, 2019; Wilmes, 2021). Multimodal interaction analysis emphasizes how an active individual performs social activities with multiple modalities. It employs low-level, high-level, and frozen activities as the study units, and analyzes the meaning and role they carry during the interaction by the modal intensity and modal complexity presented in each activity (Norris, 2004). Correlating participants' attention with modality usage, multimodal interaction analysis also proposes foreground-background continuum to distinguish participants' attention towards different activities (Norris, 2004). Moreover, it also heavily concentrates on the relationship between modal use and coordination, namely, modal structural configurations (Norris, 2004). Accordingly, multimodal interaction analysis enables us to overcome the inherent bias of "language-centeredness" and explore multimodal information transfer processes in real societies, providing a more comprehensive picture of how people use different intersectional symbols to complete social interactions. Nevertheless, the current study only involves normal people, and it remains to be further examined whether it is also applicable to special populations, such as autism.

Aiming to investigate the scientific and complete treatment mechanism of parent-child interaction and develop more effective and stable intervention methods, it is necessary to thoroughly probe the multimodal characteristics of parent-child interaction in autism.

Consequently, based on Norris'multimodal interaction analysis theory and taking individuals with ASD in five documentaries as the research objects, this study explores the multimodal communicative features in autistic parent-child interactions, thus revealing the real communicative ability of individuals with autism in parent-child interactions and bringing insights for future diagnosis, screening and intervention of autism.

2. Methods

2.1 Documentaries and Main Characters

To ensure the content of the videos analyzed were as close as possible to real parent-child interaction scenarios, five narrative documentaries were selected from a wide range of English-language autism-themed films for this study. The documentaries all truly recorded the daily performance of the main characters without any design or performance. The names and main characters of these five documentaries were Mike in *Asperger's are Us*, John in *Autism in Love*, Dina, the hero in *Dina*, Jack in *Extreme Love*, and Owen in *Life Antimated*. In this study, the quantitative data analysis included all main characters of five parent-child interaction clips, and only representative John was selected for the qualitative analysis.

2.2 Video Annotation Tools and Rules

The study employed the multimodal annotation software ELAN 6.2 (Eudico Language Annotator) to annotate the content of the documentaries. ELAN is a multimodal discourse annotation tool designed and developed by the Max Planck Institute for Psycholinguistics for creating, editing, visualizing and searching video data tags. During corpus annotation, it is capable of playing the video down to exactly 0.1 seconds and looping it if necessary.

In the present study, a certain coding scheme was developed based on Norris'nine communicative modalities and the coding scheme of Wang and Wen (2008) (see Appendix). These nine communicative modalities are spoken speech, posture, gesture, head movement, facial expression, gaze, music, scene layout, and spatial distance. In this study, the spoken language modality is divided into slow speech (SS), fast speech (SF), low speech (SL), high speech (SF), normal speech (SN), and speaking without meaning (S-N). The postures in this study were labeled as open posture (PO), closed posture (PC), forward posture (PF), backward posture (PB), and posture to complete the action (PA). The gesture modality concerns the stages and types of gesture movements, which can be specifically classified as iconic gesture (GI), metaphor gesture (GM), beat gesture (GB), deictic gesture (GD) and nonsense gesture (GN). Based on Norris'viewpoint, this study classifies head movement types into head movement (HM), head turning (HT) and head beating (HB). Facial expressions are categorized as positive facial expressions (FP) and negative facial expressions (FN). Gaze is separated into gaze at the person (GP), gaze at the central object (GO), gaze at the camera (GC) no gaze (GN). According to Hall' point of view (1966), this study distinguished four types of distance: intimate distance (DI), personal distance (DP), social distance (DS), and public distance (D-P). Music contains the sound played or sung directly by people or played out through a device. Print refers to written texts, including language, medium, layout, and content.

2.3 Research Processes

This study was divided into five main steps. In the first step, the content of parent-child interactions in the five documentaries was extracted to obtain a 40-minute analysis clip of parent-child interactions in autism. Then, the multimodal resources involved were initially recorded and counted by browsing the video content to design a coding scheme suitable for this study. The second step was to define the modality types in this study, which were divided into verbal and non-verbal modality types. The third step was to establish a layer that defines the specific patterns appearing in the study video. Based on Norris'nine communicative patterns, eight layers were established in conjunction with specific content. In the fourth step, the videos were annotated according to the coding scheme in software ELAN. The last step was the generation of the frequency and duration of each modality used by autistic individuals in parent-child interactions through the statistical function of ELAN software.

Then, based on the statistical data, the multimodal density, intensity, and structure were further analyzed and studied, and the characteristics in parent-child interactions in autism were acquired.

3. Results

3.1 Modal Distribution and Configuration in Parent-child Interaction

As shown in Table 1, the main mode mostly used by autism in parent-child interactions was the verbal mode (25.8%), followed by the postural mode (20.1%), gaze mode (14.6%), head movements (13.7%) and facial expressions (12.9%). The print mode and the music mode had the smallest proportions, 1.9% and 0.35%, respectively. It should be noted in particular that spatial distance is not included in the statistics because spatial distance in interaction is not only restricted by the interlocutor, but also by many other factors, such as the width of the table and the length of the stage. Scene layout modality was also excluded from the quantitative analysis since it exists throughout the interaction and would have an impact on the data scale if the timing of scene layout modality is included in the data.

As shown in Table 2, the total duration of normal and moderate pitch speech in the spoken mode component for individuals with autism was 369.56 seconds, representing 54.0% of the total duration of the spoken mode. The second highest proportion of speech forms was nonsense speech, with a total duration of 143.6 seconds, or approximately 21.0%.

The duration of posture mode varied considerably among individuals with autism. The duration of forward and open postures was 243.5 and 201.9 seconds, respectively, accounting for approximately 45.5% and 37.7% of the total length of the postural patterns (Table 3).

Gesture patterns were also a widely used pattern type by autistic individuals, with iconic gestures being used the longest, accounting for approximately 56.6% (Table 4). The least utilized type of gesture by them was the metaphor gesture. From the statistics, it can be concluded that autism has no significant head nodding movements during parent-child interactions. The head movement with the highest proportion and longest duration was head rotation, which occurred mainly when the autistic person's attention was drawn to the interlocutor. During the clip annotation, I found that autistic individuals showed more negative facial expressions. Statistics for the gaze pattern showed that the individuals remained gaze detached for most of the interaction time. During parent-child interactions, the musical and print modes each occurred only once and for a shorter duration.

3.2 Realization of Social Interactions in Parent-child Interaction

Through a statistical analysis of the entire parent-child interaction video, this section selects a typical case for a multimodal analysis of parent-child interactions of individuals with autism. The analysis covers not only the high-level activities of the individual with autism, but also the high-level activities of others, focusing on how the communicator achieves their goals through multiple modalities.

The clip is taken from the documentary "Extreme Love" and is located between 34:26 and 37:50 of the entire video, with a total length of 3 minutes and 24 seconds. This clip focuses on a mother who stops her autistic son from hurting himself at home. Based on the story throughout the documentary, we find out that the autistic teenager named Jack may be an elementary school student who needs to take the school bus to and from school every day. The first image appearing in this paragraph is that of his mother waiting for Jack at the school station. As we can see in Figure 1, when Jack arrives, he does not say

anything but opens his school bag and takes out his study materials for his mother. A second shot appears at home, where Jack's mother is comforting him for his mood swings and her mother also has a conversation with the researcher (Figure 3). The third shot emerging is that of the mother forcibly controlling Jack's behavior through her own body, eventually pinning Jack down.

Throughout the clip, Jack is the unconscious initiator of the interaction, and the main mode of their communication is nonverbal modality. Through the analysis of this clip, we may easily find that the highest level action of this dialogue includes the mother's greeting action, comforting action and controlling action, in addition, it contains several secondary high-level actions, such as picking up the school bag, taking off the clothes, running home, hammering the wall, knocking the leg, body pressing, and besides, there are low-level activities needed to complete the high-level activities.

The following is an analysis of modality density and modality structure configuration.

This section concerned all the modalities, like verbal language, gestures, postures, facial expressions, gaze, spatial distance, and scene layout. Scene layout refers to the non-human objects that appear throughout the interaction, including school buses, school bags, houses and cushions, which facilitate the interaction to some extent. The spatial distance between two people at the beginning of the interaction was personal distance. Then, it shifted to a more distant distance, such as social distance. After Jack and his mother met again at home, the distance became intimate distance, and then social distance and intimate distance were manifested in turn. Among all the modalities, gesture, posture and head were the three most commonly used modalities, interspersed with some speech content. Consequently, it could be summarized as follows: Firstly, this was an interaction with high modal complexity and complex modal structure; Secondly, there was an equal relationship between the modalities; Thirdly, some scene arrangement played a direct role in the interaction.

Given that this interaction involved multiple high-level activities, a foreground-background continuum was used to present the distribution and changes in participants, attention. In the first highest-level action, the interaction of picking Jack up from school, the mother followed Jack all the time, paying attention to his every movement. At first, Jack paid attention to materials such as his book bag, paper materials, and clothes, which were in his foreground. After completing a series of actions, he noticed a house not far away, so at this point, his attention shifted from the materials to the house and ran over to it, with the house in his foreground awareness and objects such as the book bag in the background awareness. In the second highest level of activity, Jack was constantly focused on himself, while his mother, consciousness shifted back and forth between the two higher levels of action of calming Jack and communicating with the researcher. In the third highest level of action, Jack was distracted three times, from the wall to his body and finally to his mother, which implied that at the beginning, the wall was in his foreground consciousness, his mother in the background consciousness, and at the end, his mother became foreground consciousness and his father and the researcher in the background consciousness. A series of high-level actions constituted Jack's foreground-background continuum. Throughout the interaction, Jack accomplished distraction from his mother's attention by putting away his book bag, carrying materials,

undressing, running, shouting, and banging on the wall. The media involved included bodies, objects, and emotions.

In summary, this clip is dominated by posture and gesture as the primary modality, and contained the interactive processes of picking up Jack from school, calming Jack, and controlling Jack's behavior as the three highest levels of activity. It had a high density of modalities and a variety of modal structure configurations. The purpose of communicator interaction was accomplished through a series of postural and gestural modalities, aided by spoken language, eye contact, and scene placement.

Table 1. Modal Distribution and Configuration in Parent-child Interaction

Mode types	Annotation number (times)	Total annotation time (second)	Annotation percentage (%)
Spoken language	174	684.38	25.8%
Posture	27	543.9	20.1%
Gesture	40	278.2	10.5%
Head movement	35	365.0	13.7%
Facial expression	34	342.3	12.9%
Gaze	62	993.3	14.6%
Music	1	9.5	0.35%
Print	1	51.7	1.9%

Table 2. Distribution and Configuration of Different Language Types in Parent-Child Interaction

Spoken language types	Annotation number (times)	Total annotation time (second)	Annotation percentage (%)
S-N	14	143.6	21.0%
SF	5	35.22	5.1%
SH	16	35.8	5.2%
SH+SS	6	10.1	1.5%

SN	97	369.56	54.0%
SS	34	76.8	11.2%
SS+SH	2	13.3	1.9%

Table 3. Distribution and Configuration of Different Posture Types in Parent-child Interaction

Posture types	Annotation number (times)	Total annotation time (second)	Annotation percentage (%)
PB	5	58.8	10.9%
PC	1	5.1	0.95%
PF	11	243.5	45.5%
PF+PO	2	25.9	4.8%
PO	8	201.9	37.7%

Table 4. Distribution and Configuration of Different Gesture Types in Parent-Child Interaction

Gesture types	Annotation number (times)	Total annotation time (second)	Annotation percentage (%)
G-N	1	21.4	7.7%
GB	5	71.7	25.8%
GD	9	26.1	9.4%
GI	24	157.5	56.6%
GM	1	1.5	0.54%

Figure 1. Jack Is Taking Some Learning Materials Out.

Figure 2. Jack Is Throwing His Clothes to His Mother.

Figure 3. Jack's Mum Is Comforting Jack

4. Discussion

4.1 Children with Autism Use Nonverbal Modalities Much More Than Verbal Modalities in Parent-child Interactions

Nonverbal modality accounted for 74.2% of the duration in parent-child interactions of individuals with autism, while the use of verbal affect accounted for only 25.8%. This suggests that nonverbal communication, such as eye contact, gestures, and joint attention, are also critical indicators of social competence (Sigman & Ruskin, 1999). Although some researchers have found that children with autism struggle with the use and comprehension of nonverbal communication (Stone et al., 1997; Maljaars et al., 2011; Buon et al., 2013; Chambon et al., 2017). Broekhofet al., (2015) have identified that children in the autism and control groups were able to understand intentional behaviors well by using an "intentional understanding task". The present study has also identified that children with ASD are capable of completing the communication process and achieving communication intentions in most cases through the cooperation between different nonverbal modalities. This study also validates previous findings that adaptive social functioning in children with autism is related to verbal and nonverbal communication, yet, it is more closely related to nonverbal communication (Kjellmer et al., 2012).

When using speech patterns, they can express themselves at a normal speed and pitch in most cases, and they are also able to express their diverse feelings by changing the speed and pitch at the appropriate time. The reason may be that most of the children in this study were high-functioning ASD children who were trained in special schools, and their communicative skills may be better than those of untrained children (Huang et al., 2018). In terms of nonverbal modality, we would notice that they spent almost

the same amount of time using different types of modality, indicating a more balanced development of their various functions. Music modality and print modality were less used in this study due to the limitations of the interactive theme and content.

4.2 Autism Uses MoreForward and Open Gestures in Parent-child Interactions

Posture serves a signaling function, typically together with body movements and facial expressions to transmit information about emotions and personality, which is of great significance in human evolution and social communication (Song et al., 2019). Among the common postures, "expansion" is an open posture in which the limbs are extended to occupy more space and is considered a dominant posture with a high sense of power; "Contracted" is a closed posture that expresses low intensity compliance when the limbs are contracted towards the trunk and occupy as little space as possible. The effectiveness difference between the two postures involves multiple levels of emotion, cognition and behavior (Song et al., 2019): Firstly, at the emotional level, researchers have found that more positive emotions are generated in open postures and more negative emotions in closed postures (Zabetipour et al., 2015); Secondly, at the cognitive level, expansive postures are perceived as more powerful than contracted postures. For example, subjects who adopt an expansive posture are more assertive in their decision making and tend to process information consistent with their decision (Fischer et al., 2011); finally, expansive postures may also implicitly activate a sense of power, which then influences an individual's fair decision making by initiating a sense of power (Li et al., 2019).

From the statistical results and video content, it could be concluded that people with autism displayed plenty of forward and open postures in parent-child interactions, like leaning toward their parents, opening their arms and hugging, etc. These series of performances indicate that people with autism are more active in expressing their thoughts, attitudes, or emotions during parent-child interactions, which contributes to their identity and promotes equitable decision-making. In addition, more open body postures suggest that they have an understanding of closer interpersonal relationships with their parents and are capable of expressing affection to their parents in specific situations. Hence, in the future, we may choose to develop deeper cognitive skills, such as identity and emotional awareness, in children with autism in parent-child interactions rather than in peer interactions or teacher- student interactions to generate better effects.

4.3 Autism Has a Lot of Head Spinning in Parent-child Interactions

Head movements convey certain communication messages in social interactions. As a form of head movement, head rotation refers to the action of keeping one's forehead and chin in the same vertical line while deviating from the direction of the trunk. It serves as one of the cues for the direction of social attention (Hietanen, 2002), while it also modulates speech wheel transitions (Duncan, 1972) and provides visual cues for speech perception (Munhall et al., 2004). It was demonstrated that children with autism showed lower head movement interpersonal synchronization in social interactions compared to typically developing children (Noel et al., 2017). Nevertheless, based on the statistical results and combined with the video content, it could be found that the autistic patients in this study showed more head rotation

during the interaction, which implies that they were able to perceive the conversation initiated by their parents and had a certain level of involvement in the interaction. Also, their head rotation would change with different topics or discourses, indicating that their attention also had a ready shift during the interaction.

In summary, individuals with autism are capable of accomplishing parent-child interactions and achieving communicative purposes through a variety of modalities. The so- called "social communication deficits" may simply represent their unique communication style, and normal people may be able to acquire these characteristics in reverse and master their communication methods to fully communicate with children with autism. In addition, modal interactions with different characteristics may also be used as an indicator to differentiate between various types or degrees of autism.

The paper examines the proportion and patterns of multimodal interactions in parent- child interactions among individuals with autism. The results of the study were as follows:

The percentage of children with autism using nonverbal modes in parent-child interactions was much higher than that ofusing verbal modes, however, the verbal mode was still the most used in terms of the nine modes. In parent-child interactions, individuals with autism have the capacity to select different modality types and structures to accomplish their communication goals depending on the interaction object, topic, and scene layout, and the selection of modalities and structures are highly individualized. Besides, in parent-child interactions, autistic children demonstrate a more positive desire to communicate and can more authentically express or reflect their emotions in a variety of ways.

Though as much of the video corpus as possible has been collected and several annotation checks have been conducted, this study still has two limitations. On the one hand, this study has not been conducted with first-hand data, which may decrease the objectivity of the study. On the other hand, there is no categorical exploration of the various subspecies of autism in this study. In future, it is recommended to investigate different subtypes of autism, record first-hand research corpus according to specific research objectives, analyze multimodal parent-child interaction characteristics of autism in more segmented scenarios, in order to offer suggestions for diagnosis, screening, and early parent-child intervention in autism.

References

Baltaxe, C. A. M. (1977). Pragmatic deficits in the language of autistic adolescents. *Journal of Pediatric Psychology*, *2*, 176-18. https://doi.org/10.1093/jpepsy/2.4.176.

Baranek, G. T. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9–12 months of age. *Journal of Autism and Developmental Disorders*, *29*, 213-224. https://doi.org/10.1023/A:1023080005650

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*, *40*(3), 293-310. https://doi.org/10.1002/dev.10034

Buon, M., Dupoux, E., Jacob, P., Chaste, P., Leboyer, M., & Zalla, T. (2013). The role of causal and intentional judgments in moral reasoning in individuals with high functioning autism. *Journal of Autism and Developmental Disorders*, 43(2), 458-470. https://doi.org/10.1007/s10803-012-1588-7

Broekhof, E., Ketelaar, L., Stockmann, L., van Zijp, A., Bos, M. G., & Rieffe, C. (2015). The understanding of intentions, desires and beliefs in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(7), 2035-2045. https://doi.org/10.1007/s10803-015-2363-3

Chambon, V., Farrer, C., Pacherie, E., Jacquet, P. O., Leboyer, M., & Zalla, T. (2017). Reduced sensitivity to social priors during action prediction in adults with autism spectrum disorders. *Cognition*, *160*, 17-26. https://doi.org/10.1016/j.cognition.2016.12.005

Chen Xianke, & Chen Liangying. (2022). Relationship between head rotation, interpersonal synchronization and autistic characteristics in autistic children. *Chinese Science Bulletin*.

De Marchena, A., & Eigsti, I. M. (2010). Conversational gestures in autism spectrum disorders: Asynchrony but not decreased frequency. *Autism Research*, *3*(6), 311-322. https://doi.org/10.1002/aur.159

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, *23*(2), 283. https://doi.org/10.1037/h0033031

El-Ghoroury, N. H., & Romanczyk, R. G. (1999). Play interactions of family members towards children with autism. *Journal of Autism and Developmental Disorders*, *29*(3), 249-258. https://doi.org/10.1023/A:1023036223397

Fine, J., Bartolucci, G., Szatmari, P., & Ginsberg, G. (1994). Cohesive discourse in pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(3), 315-329. https://doi.org/10.1007/BF02172230

Fischer, J., Fischer, P., Englich, B., Aydin, N., & Frey, D. (2011). Empower my decisions: The effects of power gestures on confirmatory information processing. *Journal of Experimental Social Psychology*, *47*(6), 1146-1154. https://doi.org/10.1016/j.jesp.2011.06.008

Grossman, R. B., Edelson, L. R., & Tager-Flusberg, H. (2013). Emotional facial and vocal expressions during story retelling by children and adolescents with high-functioning autism. *Journal of Speech Language & Hearing Research*. https://doi.org/10.1044/1092-4388(2012/12-0067)

Hall, E. T., & Hall, E. T. (1969). *The hidden dimension* (Vol. 609). Anchor.

Ham, H. S., Bartolo, A., Corley, M., Rajendran, G., Szabo, A., & Swanson, S. (2011). Exploring the relationship between gestural recognition and imitation: Evidence of dyspraxia in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*, 1-12. https://doi.org/10.1007/s10803-010-1011-1

Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, *48*, 101-119. https://doi.org/10.1016/0010-0277(93)90026-R

Haven, E. L., Manangan, C. N., Sparrow, J. K., & Wilson, B. J. (2014). The relation of parent–child interaction qualities to social skills in children wit hand without autism spectrum disorders. *Autism*, *18*(3), 292-300. https://doi.org/10.1177/1362361312470036

Hietanen, J. K. (2002). Social attention orienting integrates visual information from head and body orientation. *Psychological Research*, *66*(3), 174-179. https://doi.org/10.1007/s00426-002-0091-8

Huang, L. et al. (2018). Characteristics and intervention strategies of body language processing defects in individuals with autism. *Chinese Journal of Special Education*, *6*.

Ingersoll, B., Wainer, A. L., Berger, N. I., Pickard, K. E., & Bonter, N. (2016). Comparison of a self-directed and therapist-assisted telehealth parent-mediated intervention for children with ASD: Apilot RCT. *Journal of Autism and Developmental Disorders*, *46*(7), 2275-2284. https://doi.org/10.1007/s10803-016-2755-z

Iverson, J., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367-371. https://doi.org/10.1111/j.0956-7976.2005.01542.x

Kjellmer, L., Hedvall, A., Fernell, E., Gillberg, C., & Norrelgen, F. (2012). Language and communication skills in preschool children with autism spectrum disorders: Contribution of cognition, severity of autism symptoms, and adaptive functioning to the variability. *Research in Developmental Disabilities*, *33*(1), 172-180. https://doi.org/10.1016/j.ridd.2011.09.003

Kremer-Sadlik, T. (2004). How children with autism and Asperger Syndrome respond to questions: A 'naturalistic' theory of mind task. *Discourse Studies*, *6*(2), 185-206. https://doi.org/10.1177/1461445604041767

Kurt, O. (2011). A comparison of discrete trial teaching with and without gestures/signs in teaching receptive language skills to children with autism. *Kuram ve Uygulamada Eğitim Bilimleri*, *11*, 1436-1444.

Landa, R. J., Holman, K. C., & Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of General Psychiatry*, *64*, 853-864. https://doi.org/10.1001/archpsyc.64.7.853

Loveland, K. A., & Tunali-Kotoski, B. (2005). The school-aged child with an autism spectrum disorder. In F. R. Volkmar, R. Paul, O. N. Klin, & D. J.Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders* (pp. 247-287). (3rd ed.). New Jersey: Wiley. https://doi.org/10.1002/9780470939345.ch9

Maljaars, J., Noens, I., Jansen, R., Scholte, E., & van Berckelaer-Onnes, I. (2011). Intentional communication in nonverbal and verbal low-functioning children with autism. *Journal of Communication Disorders*, *44*(6), 601-614. https://doi.org/10.1016/j.jcomdis.2011.07.004

Mastrogiuseppe, M., Capirci, O., Cuva, S., & Venuti, P. (2015). Gestural communication in children with autism spectrum disorders during mother–child interaction. *Autism*, 19(4), 469-481. https://doi.org/10.1177/1362361314528390

Mitchell, S. et al. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics*, *27*, 69-78. https://doi.org/10.1097/00004703-200604002-00004

Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility: Head movement improves auditory speech perception. *Psychological Science*, *15*(2), 133-137. https://doi.org/10.1111/j.0963-7214.2004.01502010.x

Noel, J. P., De Niear, M. A., Lazzara, N. S., & Wallace, M. T. (2017). Uncoupling between multisensory temporal function and nonverbal turn-taking in autism spectrum disorder. *IEEE Transactions on Cognitive and Developmental Systems*, *10*(4), 973-982. https://doi.org/10.1109/TCDS.2017.2778141

Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. Routledge. https://doi.org/10.4324/9780203379493

Norris, S. (2007). The micropolitics of personal national and ethnicity identity. *Discourse & society*, *18*(5), 653-674. https://doi.org/10.1177/0957926507079633

Scudder, A., Wong, C., Ober, N., Hoffman, M., Toscolani, J., & Handen, B. L. (2019). Parent– child interaction therapy (PCIT) in young children with autism spectrum disorder. *Child & Family Behavior Therapy*, *41*(4), 201-220. https://doi.org/10.1080/07317107.2019.1659542

Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, down syndrome, and developmental delays. *Monographs of the Societyfor Research in Child Development*, *64*, 1-139. https://doi.org/10.1111/1540-5834.00005

Siller, M., & Sigman, M. (2002). The behaviors of parents of children with autism predict the subsequent development of their children's communication. *Journal of Autism and Developmental Disorders*, *32*(2), 77-89. https://doi.org/10.1023/A:1014884404276

Smith, I. M., & Bryson, S. E. (2007). Gesture imitation in autism: II. Symbolic gestures and pantomimed object use. *Cognitive Neuropsychology*, *24*, 679-700. https://doi.org/10.1080/02643290701669703

Song Lu, Zhang Qihan, Zhang Peng, et al. (2019). Psychological effects of body posture: Based on the perspective of embodiment. *Journal of Psychological Science*, *42*(4), 1004-1009.

Steger, A. (2019). Aneinander-Vorbeigehen–Eine Interaktionsanalyse flüchtiger urbaner Begegnungen. *Zeitschriftfür germanistische Linguistik*, *47*(2), 313-336. https://doi.org/10.1515/zgl-2019-0014

Stone, W. L., Ousley, O. Y., Yoder, P. J., Hogan, K. L., & Hepburn, S. L. (1997). Nonverbal communication in two-and three-year-old children with autism. *Journal of Autism and Developmental Disorders*, *27*(6), 677-696. https://doi.org/10.1023/A:1025854816091

Ticca, A. C., & Traverso, V. (2017). Participation in bilingual interactions: Translating,interpreting and mediating documents in a French social centre. *Journal of Pragmatics*, *107*, 129-146. https://doi.org/10.1016/j.pragma.2016.09.008

Wan, M. W., Green, J., & Scott, J. (2019). A systematic review of parent–infant interaction in infants at risk of autism. *Autism*, *23*(4), 811-820. https://doi.org/10.1177/1362361318777484

Wang, L. F., & Wen, Y. (2008). Using Multi-modal Approach in Applied Linguistics Research. *Technology Enhanced Foreign Language Education*, *3*, 8-12.

Wilmes, S. E., & Siry, C. (2021). Multimodal Interaction Analysis: a powerful tool for examining plurilingual students' engagement in science practices. *Research in Science Education*, *51*(1), 71-91. https://doi.org/10.1007/s11165-020-09977-z

Zabetipour, M., Pishghadam, R., & Ghonsooly, B. (2015). The impacts of open/closed body positions and postures on learners'moods. *Mediterranean Journal of Social Sciences*, *6*(2 S1), 643-643. https://doi.org/10.5901/mjss.2015.v6n2s1p643

Zhang, D. L., & Wang, Z. (2016). Theoretical Framework for Multimodal Interaction Analysis. *Foreign Languages in China*, *13*(2), 54-61.

Zlomke, K. R., & Jeter, K. (2020). Comparative effectiveness of parent–child interaction therapy for children with and without autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *50*(6), 2041-2052. https://doi.org/10.1007/s10803-019-03960-y

Appendix Coding Scheme

Modal types	Coding	Coded meaning	Coding	Coded meaning
Spoken language	SS	slow speech	SF	fast speech
	SH	high-pitched speech	SL	low-pitched speech
	SN	normal speech	S-N	nonsense speech
Posture	PO	open posture	PC	close posture
	PF	forward posture	PB	backward posture
	PA	action posture		
Gesture	GI	iconic gesture	GM	metaphor gesture
	GB	beat gesture	GD	deictic gesture
	GN	nonsense gesture		
Head movement	HM	head movement	HT	head turning
	HB	head beating		

Facial expression	FN	negative facial expression	FP	positive facial expression
Gaze	GC	gaze camera	GN	no gaze
	GO	gaze object	GP	gaze person
Music	M	music		
Print	P	print		