

## Advances and Perspectives in Research on Conversational Impairments in Autism

**Authors:** Gong Minjie, Zhang Te, Yu Wenbo, Yu Wenbo

**Date:** 2025-06-25T00:00:00+00:00

### Abstract

Social communication impairments constitute core features of autism spectrum disorders (ASD). In recent years, the atypical manifestations of this population in conversation—a prototypical social activity—have garnered focused attention from the academic community. This paper reviews the atypical manifestations of individuals with ASD across two dimensions: conversational skills and coordinated interaction (encompassing the use of filled pauses and backchanneling, prosodic entrainment, and resonance and interpersonal interaction), and proposes that future research should endeavor to address why individuals with ASD exhibit abnormal conversational performance. Regarding this issue, the author suggests that two approaches may be attempted: (1) constructing conversational competence indices and comparing differences between ASD and typically developing (TD) populations, and (2) developing cognitive computational models of conversational performance to explore differences in cognitive mechanisms between the two groups.

### Full Text

## Research Progress and Prospects on Conversational Impairments in Individuals with Autism Spectrum Disorders

**Gong Minjie<sup>1</sup>, Zhang Te<sup>2</sup>, Yu Wenbo<sup>1</sup>**

<sup>1</sup>School of Chinese Language and Literature, Nanjing Normal University, Nanjing 210097

<sup>2</sup>Fourth Ward, Harbin First Specialized Hospital, Harbin 150026

### Abstract

Social communication deficits constitute a core characteristic of autism spectrum disorders (ASD). In recent years, the atypical conversational performance of this population in typical social activities has garnered significant attention

from the academic community. This article reviews the non-typical manifestations of individuals with ASD across two dimensions: conversational skills and interactive alignment (including the use of filled pauses and backchannels, prosodic entrainment, and resonance and intersubjective engagement). We propose that future research should focus on addressing why individuals with ASD exhibit abnormal conversational performance. To this question, we suggest two potential approaches: (1) constructing conversational competence indicators and comparing differences between ASD and typically developing (TD) populations, and (2) constructing cognitive computational models of conversational performance to explore the cognitive mechanisms underlying the differences between these two groups.

**Keywords:** autism, conversational skills, interactive alignment, conversational competence, cognitive computational model

Autism spectrum disorders (ASD) are neurodevelopmental conditions that emerge during infancy, characterized by persistent deficits in social communication and restricted, repetitive patterns of behavior and interests (American Psychiatric Association 2013). Language impairment represents one of the typical features of ASD. Relevant research has predominantly adopted highly structured experimental paradigms, requiring participants to passively process experimental materials for visual (or auditory) comprehension, spoken production, or writing tasks in isolated laboratory settings. These studies detect language impairments in this population by comparing performance between ASD and TD individuals on dependent variables such as reaction time, accuracy, or neural indicators. Research has covered multiple aspects including phonetics (Chen & Peng 2021), lexical semantics (Yu Wenbo, Wang Han, Liang Dandan 2019), syntax (Dai Huilin, He Xiaowei 2021), and pragmatics (Chahboun et al. 2017). However, these studies only focus on atypical performance when ASD individuals process static linguistic symbols, neglecting the inherent social and communicative nature of language. Language serves as humanity's primary tool for communication (Fedorenko et al., 2024), and conversation represents the most important avenue for human interaction. During conversation, both parties must jointly participate and coordinate to engage in interactive activities through linguistic symbols (Fang Mei, Hu Suhong 2018). Successful conversation requires not only the correct use of static linguistic symbols but also effective processing of social information and interpersonal interaction. Do individuals with ASD experience impairments during interactions with conversational partners? What new developments have emerged in recent years? This article will first review relevant research before proposing future directions.

## 2. Conversational Skills Impairments in ASD

Conversational skills refer to a speaker's turn-taking ability (alternating between speakers), conversation initiation and maintenance ability (jointly attending to a topic), and conversation repair ability (supplementing, processing, or replac-

ing existing information) (Ninio & Snow 1996). Conversational skills impairments in ASD represent the earliest and most thoroughly studied domain. For example, Baltaxe (1977) identified three typical types of conversational skill deficits in ASD, as shown in Table 1. In the first example, the ASD individual acts as a broadcaster announcing a game rather than responding as a conversational partner. In the second example, the ASD individual violates the cooperative principle of conversation. In the third example, the pronoun “that” lacks clear reference, the interlocutors have not established common ground, and the speaker fails to provide necessary background information. Additional studies have found that ASD individuals rarely initiate topics, engage less frequently in conversation repair, and struggle to maintain conversations (De Villiers et al. 2007). Research on conversational skills in Mandarin-speaking ASD populations began later but has reached similar conclusions. For instance, studies have shown that ASD children demonstrate significantly fewer conversation initiations than TD children, employ more monotonous conversation maintenance strategies, and exhibit lower maintenance frequencies. They also show more referential impairments and lower success rates in correction (Chen Guanxing, Yang Xijie 2014; Cheng Yanhua, Ma Bosen 2022; Ma Bosen, Ni Wenjun, Zeng Xiaorong 2021; Xue Wei, Liu Min, Liu Qiaoyun 2022).

### 3. Interactive Alignment Impairments in ASD

In recent years, the academic community has increasingly focused on how ASD individuals engage in bidirectional interaction with partners during conversation, with particular attention to interactive alignment—the psychological process through which conversational partners coordinate, mutually inspire, and adapt to each other to achieve communicative goals (Jia Li, Yang Lianrui, Zhang Wenzhong 2024). Since conversation involves at least two participants, a speaker’s output largely depends on what other interlocutors have said in previous turns, and these structures in turn shape subsequent responses (Li Xiaoting 2023). Consequently, conversational content and form are interlinked, exhibiting both explicit and implicit relationships, while alignment phenomena facilitate mutual understanding and promote smooth conversation progression (Pickering & Garrod 2004). As a psychological process maintained by two (or more) parties, conversational alignment manifests across multiple aspects of conversation, which we introduce separately below.

#### 3.1.1 Filled Pauses

Natural conversation does not proceed completely fluently but contains many non-lexical structures that violate syntactic rules. Specifically, speakers unconsciously produce pauses while listeners spontaneously provide feedback. Although these structures disrupt conversational fluency, they effectively maintain conversation. The former are called filled pauses or discourse fillers, universal phenomena in speech production. Scholars widely acknowledge that filled pauses carry rich pragmatic meanings, signaling discourse production difficulties, word-

finding problems, and simultaneously conveying the speaker's intention to hold the floor (Fox Tree 2001; Li Xianyin et al. 2024). The latter are called backchannels, indicating that the listener is "attending," and though not constituting complete syntactic structures, they reflect active participation in the discourse process (Jefferson 1984; Li Xianyin et al. 2024).

In Indo-European languages, typical filled pauses include "Um" (Uhm), "Uh," and "like." Research on filled pauses in ASD populations has included children, adolescents (Irvine, Eigsti & Fein 2016; Gorman et al. 2016; McGregor & Hadden 2020; Jones, Zane & Grossman 2022; Lawley et al. 2023), adults (Lake, Humphreys & Cardy 2011; Wehrle, Grice & Vogeley 2024), and cross-linguistic comparisons (Beradze & Meir 2024). In an earlier study, Irvine, Eigsti & Fein (2016) annotated speech content from a description task, calculating two metrics: Um-rate (ratio of Um tokens to total word count) and Uh-rate (ratio of Uh tokens to total word count). Results showed significant group differences between ASD and TD participants in the former metric, which also correlated significantly with social communication questionnaire (SCQ) scores. This discriminatory effect of Um-rate between ASD and TD populations has been replicated (McGregor & Hadden 2020) and further refined. One study assessed children's structural language skills (including phonological, lexical, syntactic, and discourse abilities) and found that the number of Um tokens produced by ASD children correlated significantly with these language skills (Lawley et al. 2023). However, other research has found no obvious impairments in filler usage among ASD children. One study examined "like," which carries pragmatic meanings including: (1) approximation, (2) emphasizing subsequent content, (3) introducing the speaker's attitude or thoughts, and (4) preparing for reformulation. The study found that ASD adolescents in interview conversations showed no significant differences from TD peers in overall production frequency or subtype usage of "like," leading researchers to conclude that this population's pragmatic abilities are not universally delayed compared to TD individuals (Jones, Zane & Grossman 2022).

In studies of ASD adults, researchers asked pairs of ASD conversationalists and pairs of TD conversationalists to complete a map task, which required one participant to connect a route between start and end points based on the other's verbal description. In prosodic analyses, the most robust finding was that TD conversationalists tended to produce filled pauses with level intonation, whereas ASD conversationalists preferred rising and falling tones for discourse fillers. Additionally, both groups produced comparable numbers of fillers and favored Um over Uh, though ASD conversationalists showed greater individual variation (Wehrle, Grice & Vogeley 2024). Another study examined whether gaze patterns between conversational partners affected the frequency of speaker-oriented versus listener-oriented fillers (Clin & Kissine 2023). Researchers designed conditions with and without mutual gaze, finding that ASD adults produced significantly fewer Um-filled pauses than TD adults in both conditions, while showing opposite patterns for Uh-filled pauses: fewer in gaze conditions but more in no-gaze conditions. Similar patterns emerged for speaker-oriented fillers: ASD

adults made more breathing adjustments than TD adults in gaze conditions. This study demonstrates that eye gaze affects filler usage in ASD adults, with patterns opposite to those observed in TD populations.

Overall, most studies conclude that ASD populations exhibit abnormal filler usage compared to TD populations, though details remain controversial. Regarding age factors, participants able to engage in conversational tasks are either individuals with milder autistic traits or those who have received systematic intervention training. Consequently, they can use filled pauses to maintain conversation when acting as speakers and listeners, with obvious problems emerging only in phonetic realization and complex contexts. Regarding filler types, Um typically accompanies longer pauses, suggesting word-finding difficulties while signaling the speaker's intention to hold the floor. The pragmatic meaning of Um therefore includes turn reorganization, requiring listeners to shift conversational perspective for real-time inference. In contrast, Uh and like indicate that speakers will provide new information or explain previous content (Fox Tree 2021; Cossavella & Cevasco 2021), requiring only focused attention from listeners without involving pragmatic inference about turn organization. Thus, ASD individuals' abnormal performance with fillers like Um likely relates to their social or pragmatic impairments, while other filler types may match TD levels.

### 3.1.2 Backchannels

Unlike filled pauses, backchannels are entirely listener-oriented discourse markers, commonly taking forms such as “okay” and “En.” Research indicates that conversational partners are highly sensitive to atypical backchannel usage (e.g., frequency and phonetic realization), suggesting that backchannels are essential for conversational flow (Li 2006). Matthewman, Zane & Grossman (2022) employed a semi-structured conversational design, asking ASD adolescents to converse with research assistants (TD individuals) while recording audio, video, and eye-tracking measures. Results showed that ASD individuals provided fewer listener responses (both verbal and nonverbal) when in the listener role, though group differences in backchannels specifically did not reach significance. This suggests that ASD individuals simply provide less feedback through eye gaze and nodding during conversation. However, other studies have found that ASD conversational groups and mixed groups both produced fewer verbal and nonverbal backchannels (Rifai et al. 2022). Research on German-speaking ASD adults conducted more detailed comparisons (Wehrle, Vogeley & Grice 2024), revealing that ASD conversationalists produced fewer backchannels during initial conversation stages and showed less lexical diversity (entropy significantly lower than TD participants) in backchannel types (exactly, yes/yeah, okay, mmhm, and others). Except for mmhm, all other backchannel types also showed lower entropy in prosodic acoustic parameters, indicating more monotonous and stereotyped prosodic features.

Finally, research on Mandarin Chinese has also found that ASD adults produce

filled pauses and backchannels with prosodic features distinct from TD populations, showing higher fundamental frequency range and harmonic-to-noise ratio (Wu et al. 2024). Compared to filled pauses, research on backchannels in ASD individuals remains limited. Synthesizing these results, we conclude that ASD individuals demonstrate stronger backchannel usage abilities than filled pause abilities. This population can generally produce verbal feedback in real-time, though problems exist in richness and phonetic realization.

### 3.1.3 Summary

Although most studies identify differences between ASD and TD populations in using filled pauses and backchannels, some research finds less pronounced impairments (Vidović Zorić & Blažeković 2023). We attribute these discrepancies to differences in experimental task design and language specificity. For example, Beradze & Meir (2024) asked children to perform story-telling tasks before analyzing filled pauses, pauses, and delays, whereas studies like Jones, Zane & Grossman (2022) analyzed data from structured conversations. In story-telling tasks, research assistants serve only supportive roles, whereas in semi-structured conversations they act as conversational partners. The former involves minimal role transformation, while the latter requires constant perspective-shifting. More importantly, even among conversational tasks, significant differences exist. For instance, map tasks require no visual contact between conversationalists, relying solely on verbal communication with single-modality information, whereas other semi-structured conversations require face-to-face interaction, demanding attention to more social information (e.g., eye gaze, body movements), which poses higher demands for ASD individuals.

Furthermore, language specificity affects experimental measure selection and calculation. For example, while Um and Uh are the most common filled pauses in American English, “Eh” is most widely used in Swedish and Norwegian (Irvine, Eigsti & Fein, 2016). In Japanese, “Ne” can appear in sentence-initial, medial, and final positions with different pragmatic meanings (Yoshimura, Kawahara & Kikuchi 2020). In Mandarin Chinese, besides Zhe4ge0 and Na4ge0 commonly serving as filled pauses (Li Xianyin et al. 2024), En0 can function as both filled pause and backchannel. Wu et al. (2024) found that ASD individuals first differ from TD individuals in the phonetic realization of En0, and when En0 serves both roles, its phonetic realization shows no clear distinction, unlike in TD populations. These findings demonstrate that discourse marker usage exhibits strong language specificity and considerable flexibility, requiring differential treatment in future research.

## 3.2 Prosodic Entrainment in Conversation

During conversation, interlocutors must not only produce and perceive speech but also adjust their own production based on their partner’s performance. Prosodic entrainment describes the behavior where speakers adjust their own prosodic features after perceiving their partner’s prosodic characteristics (e.g.,

speech rate, pitch). Prosodic entrainment helps interlocutors understand each other through shared representations (Borrie, Lubold & Pon-Barry 2015), thereby ensuring smooth conversation and establishing good interpersonal relationships (Ireland 2011). Among prosodic entrainment studies on ASD populations, speech rate entrainment has been most thoroughly investigated. Wynn, Borrie & Sellers (2018) examined speech rate entrainment in TD and ASD adults and children using a perception-production quasi-conversational paradigm. Researchers first recorded audio of an English native female speaker describing pictures, then manipulated the recordings to create faster and slower versions. Participants viewed pictures while listening to audio and were asked to describe the pictures afterward. The dependent variable was participants' average speech rate (ratio of syllables produced to actual articulation time). Results showed that only TD adults demonstrated speech rate entrainment: speaking faster in fast-rate conditions and slower in slow-rate conditions, leading authors to conclude that ASD participants lacked speech rate entrainment ability.

Lehnert-LeHouillier, Terrazas & Sandoval (2020) investigated prosodic entrainment in mean fundamental frequency (F0) and F0 range in TD and ASD children using a real-time conversational task, exploring relationships with language performance, age, and nonverbal IQ. Researchers asked conversational dyads (participant and research assistant) to identify differences between their respective pictures through verbal communication. In data analysis, authors extracted the first and last thirds of conversation recordings for F0 analysis, constructed vectors from both conversationalists' acoustic measures, and calculated vector distances to determine whether prosodic entrainment occurred. They also calculated each partner's contribution rate to prosodic entrainment. Results showed that TD participants and their partners exhibited significant entrainment in mean F0, whereas ASD participants and their partners showed non-entrainment. Additionally, both groups showed negative correlations between language performance and prosodic entrainment, while nonverbal IQ showed no significant correlation with entrainment. Overall prosodic entrainment indicated that ASD participants' entrainment became more pronounced with age, whereas TD participants showed the opposite pattern, suggesting an interaction between entrainment and age. However, when considering individual contribution rates in conversation, this interaction disappeared, suggesting that conversational partners' entrainment may facilitate ASD participants' entrainment. Lehnert-LeHouillier (2022) built upon this research, finding that ASD group's F0 entrainment positively correlated with conversational efficiency and negatively correlated with turn count and total conversation duration—results inconsistent with TD control group findings. Beyond these studies, research extracting F0 contours, F0 variation trends, and rhythmic measures from both macro (sentences) and micro (syllables) discourse units has also found significant differences in prosodic entrainment between ASD and TD participants (Patel et al. 2022). Although these studies support that ASD populations show abnormal prosodic entrainment, Kruyt et al. (2023) reported that ASD participants showed no significant

abnormalities in either global prosodic entrainment measures (following Lehnert-LeHouillier & Terrazas 2020) or local prosodic entrainment measures (mean F0 in adjacent turns).

Although turn-transition duration falls outside the scope of prosodic entrainment, rapid and smooth turn transitions also result from collaborative coordination between conversational partners. Thus, “entrained” silent pauses also reflect mutual adaptation and interaction. Research has indicated that turn-transition duration reflects intimacy between conversationalists (Templeton et al. 2022). From this perspective, turn-transition duration serves a function consistent with prosodic entrainment. Wehrle et al. (2023) recruited German-speaking adult ASD and TD participants to complete a map task, measuring turn-transition times. Results found no significant group differences in turn-transition duration during conversation overall, but ASD participants showed longer turn-transition times in early conversation stages. Since transition time inversely correlates with conversational efficiency and interlocutor intimacy, this result suggests that ASD populations’ adaptation and coordination abilities in conversation are slightly weaker than those of TD populations.

Overall, ASD populations demonstrate less prosodic entrainment than TD populations (Kruyt & Beňuš 2021). However, current research remains in its early stages and has not involved tonal languages. Since Chinese employs lexical tones, F0 calculation must consider both intonation and tone, making the situation more complex. Additionally, as Chinese syllables have relatively consistent duration (approximately 250ms), conversational partners may not show obvious entrainment or disentrainment in speech rate.

### 3.3 Resonance and Intersubjective Interaction in Conversation

Dialogic syntax, proposed by Bois (2014), provides a theoretical framework for describing interactive engagement between conversational partners, analyzing resonance in form and meaning, and exploring structural coupling mechanisms between utterances (Wang Deliang 2018, 2021). Resonance represents the core concept, referring to how one utterance reproduces partial structures or content from a previous utterance, serving as a catalyst for intimacy between discourse units (Bois 2014), as shown in Example (1) (Wang Deliang 2021):

- (1) A: Oh my.  
B: Are you busy now?  
A: Whether I’m busy depends on myself. I’ve been okay these two days.  
B: All up to your own arrangement.

Speaker A’s response resonates with Speaker B’s question through the structure “紧不紧” (whether busy), while Speaker B’s response resonates with Speaker A’s utterance through “自己” (self). These two instances constitute static resonance, where speakers simply repeat their partner’s linguistic structures. Another type involves speakers creatively reworking previous structures, termed dynamic resonance, which demonstrates higher intersubjective engagement and plays an

important role in child language acquisition (Tantucci & Wang 2022). In Example (2) (Tantucci & Wang 2022:556), the child produces two instances of creative resonance: first negating the mother's utterance “是” (is), then creatively reworking their own utterance “在给” (giving) through “也” (also) to emphasize subjective intention.

- (2) Child: I'm giving the calf a bath.  
Mother: She is a guest.  
Child: Not a guest.  
Child: I also want to give the calf a bath.

Drawing on dialogic syntax theory, some studies have analyzed corpus data, finding that ASD children possess resonance abilities but produce more atypical resonance than TD children, as in Example (3) (Hobson et al. 2012):

- (3) Examiner: What do you like most about yourself?  
Child: Most about myself is the teach.

In research on Mandarin-speaking ASD children, Tantucci & Wang (2023) analyzed nearly 20,000 turns of conversation, focusing on relationships between static resonance, dynamic resonance, and sentence final particles (SFPs). Results showed that ASD children tended to produce dynamic resonance on their own utterance content, whereas TD children more frequently reworked others' utterances. More importantly, among two linguistic structures reflecting intersubjective engagement (SFPs and dynamic resonance), ASD children tended to use only one type, while TD children often used both simultaneously, suggesting that competition between these two engagement forms may underlie ASD children's weaker resonance abilities. Additionally, Wang Deliang, Jiang Yuanqun (2022) analyzed conversational data from five-year-old Mandarin-speaking ASD children, finding no significant difference in resonance quantity compared to TD children but substantial individual variation. The dialogic syntax perspective, similar to prosodic entrainment research, focuses on interactive relationships between conversationalists and emphasizes connections between language, cognition, and intersubjective engagement, aligning well with the social communication deficits characteristic of ASD. However, dialogic syntax research predominantly employs corpus analysis methods, lacking exploration of real-time data and paying limited attention to phonetic and non-verbal information.

#### 4. Other Studies on Conversational Impairments in ASD

The aforementioned studies all evaluate ASD individuals' conversational performance from the perspective of experimenters or third parties, lacking direct input from conversational partners. Sng et al. (2020) collected 258 questionnaires from conversational partners of ASD individuals, primarily parents and friends around 30 years old. The questionnaire asked participants to rate ASD individuals' conversational performance on a three-point scale, including items such as “inappropriately initiates conversation,” “does not make eye contact during communication,” and “fails to explain or clarify when explicitly indicating

misunderstanding.” Results showed that conversational partners most frequently reported impairments in “inappropriately initiating conversation,” “mistakenly assuming conversational partners have prior knowledge,” “inability to express emotions,” and “narrative difficulties.” These findings align generally with commonly recognized ASD conversational impairments, but conversational partners did not identify some repeatedly mentioned impairment types in previous research, such as “echolalia” and “overly detailed verbal descriptions.” This result indicates discrepancies between researchers’ assessments and conversational partners’ subjective impressions, with many researcher-identified impairments not recognized by conversational partners. Larkin et al. (2017) similarly employed an ecologically valid paradigm, asking two raters to evaluate conversational performance in ASD children and language-ability-matched children, rating appropriate production of filled pauses and backchannels, necessary conversational repairs, and appropriate gaze. Results showed ASD participants significantly lagged behind controls on multiple measures, with conversational ratings significantly correlating with pragmatic ability scores.

With the popularization of neuroscience techniques, some scholars have explored the neural mechanisms underlying ASD conversational impairments. Related studies predominantly employ hyper-scanning technology, requiring two participants to complete conversational tasks while recording neural signals. Results have found that German-speaking ASD adults show weaker neural synchronization in the temporoparietal junction (TPJ) compared to TD participants (Quinoñes-Camacho et al. 2021). This brain region is closely related to social information processing and mental state inference, with signal strength also negatively correlating with social communication impairment severity.

## 5. Future Research Directions

The above review demonstrates substantial research on conversational impairments in ASD individuals, encompassing both unilateral descriptions of autistic individuals (e.g., conversational skills) and summaries of interactive processes (e.g., prosodic entrainment). However, current research has yet to address the most fundamental question: Why do ASD populations exhibit atypical conversational performance? What are the root causes of conversational skill deficits and alignment impairments? To address this question, we propose exploration along two paths.

### 5.1 Extracting Conversational Competence Indicators for Both Populations

The reviewed articles reveal clear differences between ASD and TD populations from various perspectives and methods, but these differences primarily reflect conversational performance rather than conversational competence. We propose that conversational competence refers to relatively stable underlying traits exhibited during conversation that enable individuals to effectively initiate, smoothly maintain, and appropriately terminate conversations. The con-

versational skills introduced in this article (e.g., conversation repair), use of backchannels and filled pauses, and prosodic entrainment between conversationalists all represent external manifestations of conversational competence. Theoretically, conversational competence should represent a stable trait across populations. Even though ASD populations show obvious impairments in conversational performance, the constituent dimensions of their conversational competence should align with those of TD populations, with differences appearing in specific manifestations of each dimension and interactions between dimensions.

Regarding indicator extraction, conversational competence should primarily consider dyadic conversation as the main context, with verbal performance as core content and cognitive performance as supplementary content. Verbal performance should incorporate information from phonetic, lexical, semantic, and syntactic levels, while cognitive performance should consider attention allocation and working memory capacity. Finally, we must carefully consider the relationship between conversational competence and Chomsky's notion of linguistic competence. Chomsky's narrow linguistic competence includes only the ability to use linguistic symbols (particularly recursive ability), while broad linguistic competence overlaps to some extent with conversational competence (Hauser, Chomsky & Fitch 2002). Undeniably, conversation involves both processing linguistic symbols and mutual coordination between conversationalists. When listeners fail to process speakers' linguistic symbols (e.g., due to noise interference, syntactic ambiguity, or unfamiliar low-frequency vocabulary), they prompt speakers to explain and clarify through backchannels. Simultaneously, speakers must monitor listeners' reactions during production and adjust their speech through filled pauses. Therefore, we propose that conversational competence should include both linguistic competence and interactive alignment ability, with the former having relatively smaller weight. This approach can encompass all elements necessary for maintaining conversation while addressing different ASD subgroups.

## 5.2 Constructing Cognitive Computational Models of Atypical Conversational Performance in ASD

In recent years, using cognitive computational modeling to investigate human psychological and linguistic performance has become a trend. When conducting cognitive modeling research, the first step involves establishing a basic cognitive framework or model. Currently, predictive coding theories (PCT) are widely accepted, positing that the brain's fundamental operation involves predicting future events, with ASD impairments stemming from an inability to establish or update internal beliefs (models) for future prediction (Jing Wei et al. 2024; Cui Ke, Luo Fei, Wang Jinyan 2024; Sinha et al. 2014). PCT includes several core concepts: prior beliefs refer to established internal beliefs, likelihood refers to the probability of external information occurring under current internal beliefs, and posterior beliefs represent the weighted value between the two—i.e., internal belief updating after external information occurs. According to

Bayes' formula, if prior beliefs have concentrated probability distributions that align with likelihood, posterior beliefs become more stable, more informative, and more predictive of the future. As a social behavior, conversation involves high uncertainty. Without relying on prior beliefs for prediction, the brain would need to continuously process input information (speech, eye gaze, facial expressions, and movements), consuming enormous cognitive resources. Therefore, individuals must accumulate experience to form reasonable prior beliefs and rapidly adjust posterior beliefs based on current contexts for subsequent predictions.

Regarding ASD populations' atypical conversational performance, two explanations are possible: (1) Individuals cannot form reasonable prior beliefs through life experience, i.e., they cannot learn basic conversational patterns through internal or external reinforcement; (2) Individuals possess correct prior beliefs but cannot update posterior beliefs in a timely manner, and due to inadequate regulation, cannot coordinate large discrepancies between predicted and actual values. We propose that future research can validate these two explanations using reinforcement learning and active inference (Cannon et al. 2021), respectively, to unravel the puzzle of ASD populations' atypical conversational performance.

The approach of extracting conversational competence indicators is proposed from an interdisciplinary perspective of linguistics and psychology, while constructing cognitive computational models of conversational performance relies on deep integration between linguistics and computer science. The former emphasizes conversational competence as an underlying trait, which in rehabilitation training can avoid piecemeal interventions and prevent children from demonstrating false intervention effects through guessing and memorization. The latter attempts to quantify and visualize conversational processes, which not only helps us understand basic cognitive processes underlying conversation but also serves refined intervention protocols.

## References

- Chen Guanxing, Yang Xijie. (2014). An Exploration of Conversational Competence in Children with Autism. *Chinese Journal of Special Education*, (11), 45-52.
- Cheng Yanhua, Ma Bosen. (2022). A Comparative Study of Conversational Repair Behaviors Between High-Functioning Autistic Children and Typically Developing Children. *Chinese Journal of Special Education*, (02), 28-36+44.
- Cui Ke, Luo Fei, Wang Jinyan. (2024). Sensory Perception Abnormalities in Autism Under the Bayesian Framework. *Chinese Science Bulletin*, 69(Z1), 489-498.
- Dai Huilin, He Xiaowei. (2021). Comprehension of Chinese Wh-Questions by Children with Specific Language Impairment and High-Functioning Autism. *Modern Foreign Languages*, 44(04), 522-535.

- Fang Mei, Hu Suhong. (2023). Interactive Linguistics from a Cross-Linguistic Perspective: A Case Study of Syntactic Co-Construction. *Contemporary Linguistics*, 25(02), 159-178.
- Jia Li, Yang Lianrui, Zhang Wenzhong. (2024). The Influence of L2 Pragmatic Competence and Working Memory on L2 Pragmatic Alignment Effects. *Foreign Language Teaching and Research*, 56(06), 854-866+959.
- Jing Wei, Chen Qi, Xue Yunqing, Yang Miao, Zhang Jie. (2024). Predictive Coding Deficits in Individuals with Autism: Abnormal Feedforward or Feedback Connections? *Advances in Psychological Science*, 32(05), 813-833.
- Li Xianyin, Zhang Wenxian. (2023). *Interactive Grammar of Spoken Chinese: Observations Based on Time Management*. China Social Sciences Press.
- Li Xiaoting. (2023). *Multimodal Interaction and Turn-Taking in Chinese Conversation* (Translated by Yao Shuangyun et al.). China Social Sciences Press.
- Ma Bosen, Ni Wenjun, Zeng Xiaorong. (2021). A Comparative Study of Other-Initiated Self-Repair Strategies Between Autistic Children and Typically Developing Children. *Language Strategic Research*, 6(06), 23-32.
- Wang Deliang, Jiang Yuanqun. (2022). A Study of Conversations of Mandarin-Speaking Autistic Children Based on Dialogic Syntax. *Journal of Tianjin Foreign Studies University*, 29(04), 87-101+113.
- Wang Deliang. (2018). On Resonance in Dialogic Syntax. *Contemporary Linguistics*, 20(02), 214-229.
- Wang Deliang. (2021). On the “Syntax” of Dialogic Syntax. *Contemporary Linguistics*, 23(01), 97-113.
- Xue Wei, Liu Min, Liu Qiaoyun. (2022). Characteristics of Conversational Competence in Children with Autism in Different Conversational Contexts. *Journal of Shaanxi Xueqian Normal University*, 38(10), 41-48.
- Yu Wenbo, Wang Han, Liang Dandan. (2019). The Influence of Verb Implicit Causality on Pronoun Processing in Mandarin-Speaking High-Functioning Autistic Children Aged 5-7. *Psychological Science*, 42(03), 577-583.
- Baltaxe, C. A. M. (1977). Pragmatic Deficits in the Language of Autistic Adolescents. *Journal of Pediatric Psychology*, 2(4), 176-180.
- Beradze, M., & Meir, N. (2024). Disfluencies as a Window into Pragmatic Skills in Russian-Hebrew Bilingual Autistic and Non-Autistic Children. *Journal of Autism and Developmental Disorders*.
- Bois, J. W. D. (2014). Towards a dialogic syntax. *Cognitive Linguistics*, 25(3), 359-410.
- Borrie, S. A., Lubold, N., & Pon-Barry, H. (2015). Disordered speech disrupts conversational entrainment: A study of acoustic-prosodic entrainment and com-

municative success in populations with communication challenges. *Frontiers in Psychology*, 6.

Cannon, J., O' Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence. *Autism Research*, 14(4), 604-630.

Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2017). Can you tell it by the prime? A study of metaphorical priming in high-functioning autism in comparison with matched controls. *International Journal of Language & Communication Disorders*, 52(6), 766-785.

Chen, F., & Peng, G. (2021). Categorical Perception of Pitch Contours and Voice Onset Time in Mandarin-Speaking Adolescents With Autism Spectrum Disorders. *Journal of Speech, Language, and Hearing Research*, 64(11), 4436-4451.

Cossavella, F., & Cervasco, J. (2021). The importance of studying the role of filled pauses in the construction of a coherent representation of spontaneous spoken discourse. *Journal of Cognitive Psychology*, 33(2), 172-186.

De Villiers, J., Fine, J., Ginsberg, G., Vaccarella, L., & Szatmari, P. (2007). Brief Report: A Scale for Rating Conversational Impairment in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 37(7), 1375-1380.

Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. F. (2024). Language is primarily a tool for communication rather than thought. *Nature*, 630(8017), 575-586.

Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, 29(2), 320-326.

Gorman, K., Olson, L., Hill, A. P., Lunsford, R., Heeman, P. A., & Van Santen, J. P. H. (2016). Uh and um in children with autism spectrum disorders or language impairment. *Autism Research*, 9(8), 854-865.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298(5598), 1569-1579.

Hobson, R. P., Hobson, J. A., García-Pérez, R., & Du Bois, J. (2012). Dialogic Linkage and Resonance in Autism. *Journal of Autism and Developmental Disorders*, 42(12), 2718-2728.

Irvine, C. A., Eigsti, I.-M., & Fein, D. A. (2016). Uh, Um, and Autism: Filler Disfluencies as Pragmatic Markers in Adolescents with Optimal Outcomes from Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 46(3), 1061-1070.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens "Yeah" and "Mm Hm". *Paper in Linguistics*, 17(2), 197-216.

- Jones, R., Zane, E. R., & Grossman, R. B. (2022). Like, it' s important: The frequency and use of the discourse marker like in older autistic children. *Autism & Developmental Language Impairments*, 7, 23969415221129132.
- Kruyt, J., & Beňuš, Š. (2021). Prosodic entrainment in individuals with autism spectrum disorder. *Topics in Linguistics*, 22(2), 47-61.
- Kruyt, J., Polónyiová, K., Ostatníková, D., & Beňuš, Š. (2023). Global and local prosodic entrainment in a task-oriented interaction in autistic and neurotypical children. In M. Alikhani, J. Parish-Morris, E. Prud' hommeaux, V. Romero, & T. Chowdhury (Eds.), *Proceedings of the First Workshop on Connecting Multiple Disciplines to AI Techniques in Interaction-centric Autism Research and Diagnosis (ICARD 2023)* (pp. 1-11). Association for Computational Linguistics.
- Lake, J. K., Humphreys, K. R., & Cardy, S. (2011). Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders. *Psychonomic Bulletin & Review*, 18(1), 135-140.
- Larkin, F., Hobson, J. A., Hobson, R. P., & Tolmie, A. (2017). Collaborative competence in dialogue: Pragmatic language impairment as a window onto the psychopathology of autism. *Research in Autism Spectrum Disorders*, 43-44, 14-24.
- Lawley, G. O., Bedrick, S., MacFarlane, H., Dolata, J. K., Salem, A. C., & Fombonne, E. (2023). "Um" and "Uh" Usage Patterns in Children with Autism: Associations with Measures of Structural and Pragmatic Language Ability. *Journal of Autism and Developmental Disorders*, 53(8), 2986-2997.
- Lehnert-LeHouillier, H., Terrazas, S., & Sandoval, S. (2020). Prosodic Entrainment in Conversations of Verbal Children and Teens on the Autism Spectrum. *Frontiers in Psychology*, 11.
- Matthewman, H., Zane, E., & Grossman, R. (2022). Comparing Frequency of Listener Responses Between Adolescents with and Without ASD During Conversation. *Journal of Autism and Developmental Disorders*, 52(3), 1007-1018.
- McGregor, K. K., & Hadden, R. R. (2020). Brief Report: "Um" Fillers Distinguish Children With and Without ASD. *Journal of Autism and Developmental Disorders*, 50(5), 1816-1821.
- Ninio, A., & Snow, C. (1996). *Pragmatic Development* (1st ed.). Routledge.
- Patel, S. P., Cole, J., Lau, J. C. Y., Fragnito, G., & Losh, M. (2022). Verbal entrainment in autism spectrum disorder and first-degree relatives. *Scientific Reports*, 12(1), 11496.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2), 169-190.
- Quiñones-Camacho, L. E., Fishburn, F. A., Belardi, K., Williams, D. L., Hupert, T. J., & Perlman, S. B. (2021). Dysfunction in interpersonal neural syn-

chronization as a mechanism for social impairment in autism spectrum disorder. *Autism Research*, 14(8), 1585-1596.

Rifai, O. M., Fletcher-Watson, S., Jiménez-Sánchez, L., & Crompton, C. J. (2022). Investigating Markers of Rapport in Autistic and Nonautistic Interactions. *Autism in Adulthood*, 4(1), 3-11.

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., Diamond, S. P., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences*, 111(42), 15220-15225.

Sng, C. Y., Carter, M., Stephenson, J., & Sweller, N. (2020). Partner Perceptions of Conversations with Individuals with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 50(4), 1182-1197.

Tantucci, V., & Wang, A. (2022). Dynamic Resonance and Explicit Dialogic Engagement in Mandarin First Language Acquisition. *Discourse Processes*, 59(7), 553-574.

Tantucci, V., & Wang, A. (2023). Dialogic Priming and Dynamic Resonance in Autism: Creativity Competing with Engagement in Chinese Children with ASD. *Journal of Autism and Developmental Disorders*, 53(6), 2458-2474.

Templeton, E. M., Chang, L. J., Reynolds, E. A., Cone LeBeaumont, M. D., & Wheatley, T. (2022). Fast response times signal social connection in conversation. *Proceedings of the National Academy of Sciences*, 119(4), e2116915119.

Vidović Zorić, A., & S. Blažeković. (2023). Pragmatic function of speech disfluencies in high-functioning children with autism spectrum disorder. *Govor*, 40(2), 147-170.

Wehrle, S., Cangemi, F., Janz, A., Vogeley, K., & Grice, M. (2023). Turn-timing in conversations between autistic adults: Typical short-gap transitions are preferred, but not achieved instantly. *PLOS ONE*, 18(4), e0284029.

Wehrle, S., Grice, M., & Vogeley, K. (2024). Filled Pauses Produced by Autistic Adults Differ in Prosodic Realisation, but not Rate or Lexical Type. *Journal of Autism and Developmental Disorders*, 54(7), 2513-2525.

Wehrle, S., Vogeley, K., & Grice, M. (2024). Backchannels in conversations between autistic adults are less frequent and less diverse prosodically and lexically. *Language and Cognition*, 16(1), 108-133.

Wu, V. S.-H., Liu, Y. H., Lyu, S., Pan, H., & Gau, S. S.-F. (2024). Backchannel and Filled Pause by Taiwan Mandarin Speakers with Autism. *5th International Symposium on Applied Phonetics (ISAPh 2024)*, 131-136.

Yoshimura, Y., Kawahara, K., & Kikuchi, M. (2020). Turn-taking in Children with Autism Spectrum Disorder: Discussion regarding Ne and Backchannel Interjections. *Frontiers in Psychology*, 11, 1234.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv – Machine translation. Verify with original.*