Alignment of Representational Complexity as a Latent Control Parameter in Referential Communication

Amir-Homayun Hallajian (Amirhomayun.Hallajian@donders.ru.nl)
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, the Netherlands

Saskia B. J. Koch (Saskia.Koch@donders.ru.nl)
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, the Netherlands

Jana Bašnáková (Jana.Basakanova@donders.ru.nl)
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, the Netherlands

Ivan Toni (Ivan.Toni@donders.ru.nl)*
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, the Netherlands

Arjen Stolk (Arjen.Stolk@dartmouth.edu)*
Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA
*These authors contributed equally

Abstract

Despite large inter-individual differences in experience and conceptual structures, humans converge on referents largely underspecified by the signals exchanged during communication. Many neural networks have become extremely sensitive to context-dependent relationships between signals, but remain relatively blind to their referents outside the signal space. Here, we study how human interlocutors dynamically coordinate both their signal and referential spaces over extended communicative interactions. We identify a latent control parameter, representational complexity, that may regulate referential coordination in human communication. Using a custom hierarchical Transformer model, we generate movement- and interaction-level embeddings from neurotypical (NT) and autistic (ASC) dyads engaged in an experimental semiotic task. By leveraging ASC-related communicative variance, we identify changes in parameters that track the representational dimensionality of signals and referents as communication unfolds. Movement-level embeddings (within-trial dependencies) could not differentiate the two groups, indicating comparable communicative behaviors. In contrast, interaction-level embeddings (across-trial dependencies) distinguished ASC from NT dyads with high accuracy. Crucially, representational complexity, i.e. dyadic alignment in the interaction-level intrinsic dimensionality used to encode communicative histories, tracked referential coordination demands, with greater misalignment in ASC dyads under referential volatility. These findings suggest that referential alignment is an interaction-level process, driven by the dynamic adaptation of representational complexity, rather than statistical relationships between signals alone.

Keywords: Conceptual alignment; Social interaction; Autism spectrum condition; Natural language processing

Introduction

Human communication is often conceptualized as signal transmission, presupposing that interlocutors encode and decode signals within a shared lexicon (Gaskell & Altmann, 2007). Neural network models extend this principle by predicting linguistic item occurrence based on statistical patterns in large text corpora, e.g., (Contreras Kallens et al., 2023; Devlin et al., 2018; Lepori, 2020). However, real-world dialogue relies on signals whose referential content is highly contingent on evolving local contexts, often requiring interlocutors to generate novel signal-referent mappings beyond what purely distributional models capture (Reddy, 1993; Stolk et al., 2016). While theoretical accounts describe the complexities of referential communication (Christiansen & Chater, 2022; Clark, 1996; Levinson, 2020), it remains unclear how interlocutors regulate the referential process to effectively coordinate novel, context-dependent mappings in real-time interactions.

To address this, we combine two complementary approaches to characterize interaction-specific mappings between signals and referents across communicative turns. First, we employ the Tacit Communication Game (TCG), an experimental semiotic task that amplifies natural generative demands by requiring participants to communicate without preexisting shared signal-referent mappings (Galantucci & Garrod, 2011; Stolk et al., 2013). In the TCG, dyads collaborate to arrange geometric shapes into designated configurations across multiple turns, minimizing reliance on conventional linguistic or gestural cues while enabling precise quantification of communicative behaviors across diverse referential challenges. Second, we develop a hierarchical Transformer model (HTM) (Vaswani et al., 2017) to generate movement- and interaction-level embeddings of communicative behaviors in 113 adult dyads. Unlike standard large language models, this approach captures dependencies not only between tokens, but also between the actual referents of those tokens over the broader communicative exchange. By leveraging full access to both signal trajectories and referential spaces, we move beyond surface-level signal analysis to uncover how discrete behavioral sequences evolve into structured patterns of referential coordination over time.

By training the Transformer on dyads from a neurotypical (NT) population and on dyads from linguistically matched

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

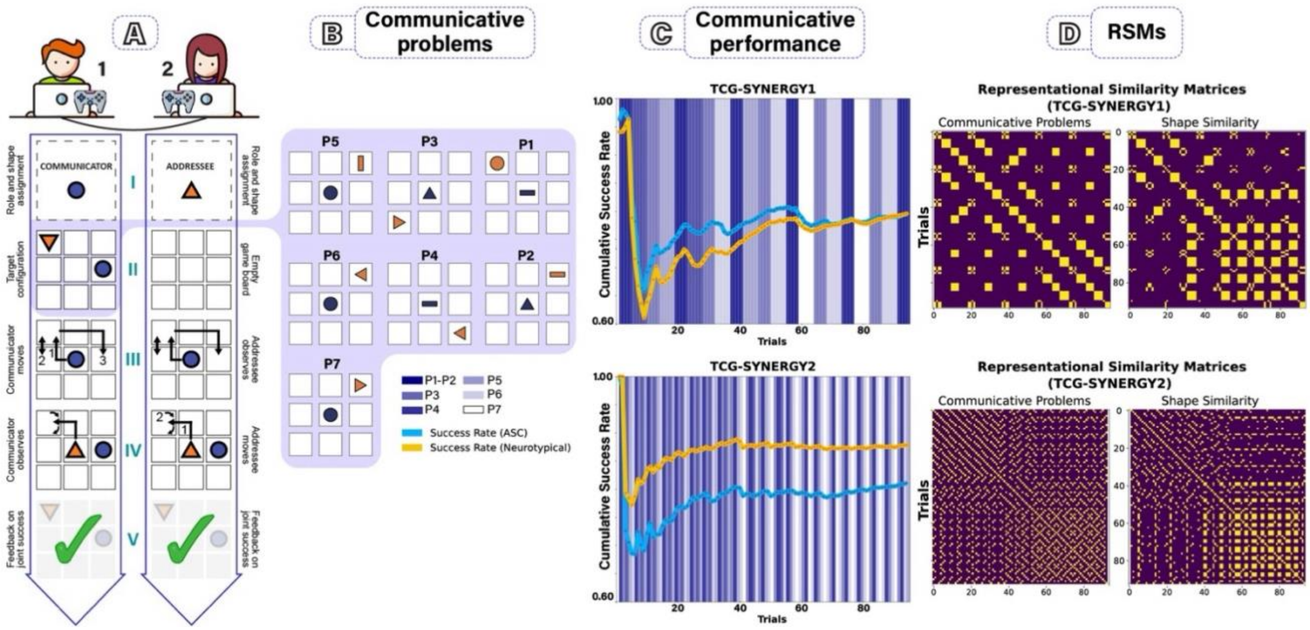

Figure 1: Task design, problem categories, and referential coordination. **(A)** In the TCG, dyads alternate between Communicator and Addressee roles on subsequent trials (one shown here), working together to reproduce target configurations of their assigned shapes on a digital game board. Since the Addressee lacks direct access to the target, the Communicator must convey the required position of the Addressee's shape through stepwise movements of their own shape (epoch II). Success requires continuous coordination, as each trial presents novel communicative problems. **(B)** Task demands vary based on token assignments and target configurations. For instance, when geometric incompatibilities prevent direct matching, dyads must develop signals beyond simple matching strategies to communicate the Addressee's target location and orientation. **(C)** Cumulative success rates indicate that NT (orange) and ASC (blue) dyads performed comparably in TCG-SYNERGY1, where problems followed a structured sequence. However, ASC dyads faced greater difficulty in TCG-SYNERGY2, where the unpredictable problem order increased referential coordination demands. **(D)** Hypotheses-driven RSMs capture referentially relevant similarity between trials, including problem and shape similarity across target configurations.

individuals on the autism spectrum (ASC), we leverage ASC-related communicative challenges to identify latent control parameters governing referential coordination in dyadic interaction (Lord et al., 2020; Wadge et al., 2019). More precisely, we examine parameters extracted from high-dimensional projections of observed communicative behaviors, such as the intrinsic dimensionality used by a dyad to encode interaction histories (Jazayeri & Ostojic, 2021; Roads & Love, 2024), and assess whether these latent metrics of representational complexity predict variability in communicative success. Additionally, by training and testing on independent datasets, we evaluate the generalizability of these control parameters across diverse communicative contexts, shedding light on the computational mechanisms that govern human referential communication.

Methods

Communication task and datasets

This study analyzed behavioral data from two independent datasets, TCG-SYNERGY1 (60 dyads: 34 NT and 26 ASC) and TCG-SYNERGY2 (53 dyads: 28 NT and 25 ASC) collected at the Donders Centre for Cognitive Neuroimaging. Both datasets use the TCG, a real-time interactive task in which dyads coordinate to reproduce target configurations of geometric shapes on a 3x3 digital game board displayed on separate screens (Figure 1A). Each dyad completed 94 trials, alternating roles between Communicator and Addressee. The Communicator viewed the target configuration and conveyed the required position of the Addressee's shape solely through stepwise movements of their own shape (horizontal/vertical translations and 90° rotations) via a game controller. Lacking direct access to the target configuration, the Addressee interpreted the Communicator's movements to position their shape accordingly. Success required continuous coordination within a dyad, as each trial presented novel communicative problems with varying shape constraints (Figure 1B). For instance, some trials introduced incompatible shape orientations, forcing dyads to develop signals beyond simple matching strategies. While TCG-SYNERGY1 presented communicative problems in structured blocks, TCG-

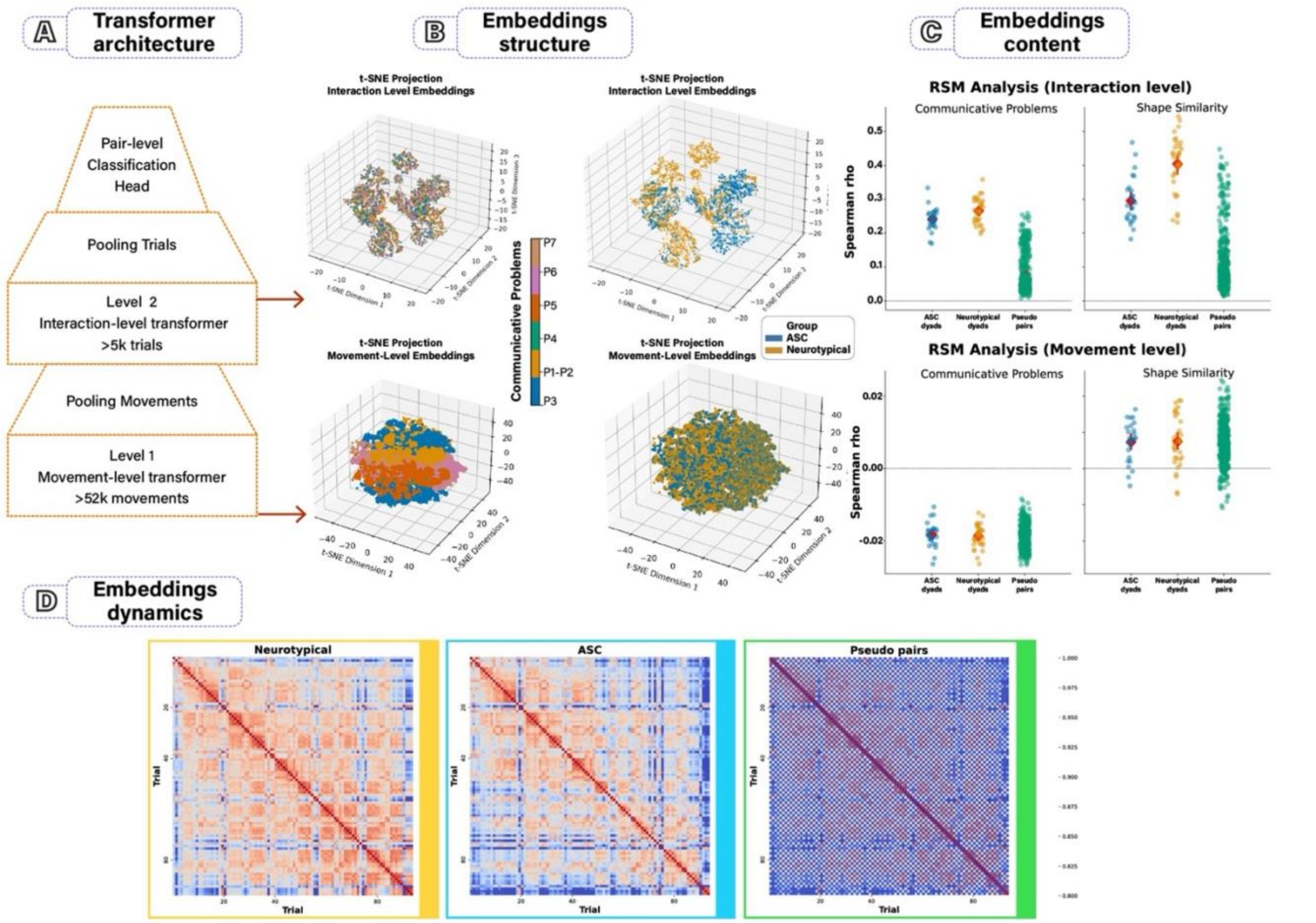

Figure 2: Hierarchical Transformer model and encoding of referential structure. **(A)** The HTM generates movement- and interaction-level embeddings from dyads' communicative behaviors. **(B)** t-SNE projections of these embeddings reveal distinct representational structures. Movement-level embeddings (bottom) cluster by signal properties (left) but fail to differentiate groups (right). Interaction-level embeddings (top) distinguish ASC from NT dyads, preserving within-dyad signal organization. **(C)** RSM analyses show that interaction-level embeddings (top), but not movement-level embeddings (bottom), capture referential task structure, with NT dyads aligning more strongly with task constraints than ASC dyads and pseudo pairs. **(D)** NT dyads exhibit stronger long-range dependencies in their interaction-level RSMs than ASC dyads.

SYNERGY2 introduced them in an unpredictable order, increasing referential volatility and disproportionally impacting ASC dyads' communicative performance (Figure 1C – the performance difference could not be due to AQ scores differences between ASC groups, $t_{(48.98)}$=0.06, p>.05). To assess model generalization across communicative contexts, we trained the Transformer model on TCG-SYNERGY1 (52,708 movements) and tested it on TCG-SYNERGY2 (53,143 movements).

Model architecture and training framework

We developed a HTM to capture movement-level (within-trial) and interaction-level (across-trial) dependencies in communicative behavior. The model consists of two processing levels: a movement-level Transformer that extracts fine-grained movement dynamics within each trial, and an interaction-level Transformer that models how trial-by-trial dependencies evolve across dyadic interactions. Both transformers share the same architecture, each with 256 embedding dimensions, 8 attention heads, 6 layers, and a dropout rate of 0.4. To generate summary embeddings for binary-classification of NT and ASC dyads, we employed an attention-based pooling mechanism followed by a linear classification head (Figure 2A). For training, we used the Adam optimizer with a learning rate of 2×10^{-5}, adjusted via a cosine warm-up scheduler to ensure stable convergence.

Input features. The model takes as input Communicator shape movement sequences, represented by *x-y coordinates*, *orientation*, and *normalized time*, which capture spatial and temporal aspects of movement. In addition, Addressee target-encoding variables provide task-relevant information, including a *closeness-score* (Manhattan distance to the target position), *reached-target* (binary indicator of target location reached), *angle-match* (binary indicator of target orientation matched), and *shapes-geometry* (encoded representation of the relative geometric compatibility between the Communicator's and Addressee's shape). Sequences were padded to the longest trial, with padding tokens excluded from self-attention computations.

Movement-level embeddings. Movement sequences were projected into a high-dimensional latent space using a fully connected embedding layer with positional encoding. A multi-layer Transformer encoder then captured movement dependencies, generating movement-level embeddings via self-attention. An attention-based pooling mechanism condensed these embeddings into trial-level vector representations.

Interaction-level embeddings. Trial-level embeddings were inputs into the interaction-level Transformer, which modeled trial-by-trial dependencies within dyadic interactions. A dyad-aware masking mechanism restricted self-attention operations to dependencies within each dyad. A second pooling mechanism then aggregated trial-level outputs into dyad-level representations, capturing interaction-wide patterns.

Classification and training. The classification head comprised a linear layer with dropout regularization, producing logits for binary NT vs. ASC classification. The model was trained in PyTorch Lightning (Paszke et al., 2019) with GPU acceleration, using a combination of supervised contrastive loss (Khosla et al., 2020) to enhance intra-class clustering and inter-class separation, and binary cross-entropy loss to optimize classification accuracy. The final loss function balanced both components to optimize representation learning and classification performance.

Model performance evaluation

Classification performance. To assess model performance, embeddings were extracted from the test set (TCG-SYNERGY2) and evaluated across hierarchical levels. Logistic regression was applied to embeddings, with different cross-validation strategies. At the dyad level, Leave-One-Out Cross-Validation was employed due to the limited number of dyads and the need for unbiased evaluation of each dyad. At the movement- and interaction-level, standard cross-validation was utilized, leveraging the larger dataset sizes at these levels.

Representational similarity analysis (RSA). To examine the model's representational structure, we constructed trial-by-trial representational similarity matrices (RSMs) by computing pairwise cosine similarities between empirical embeddings. Empirical RSMs were correlated with hypothesis-driven RSMs (Figure 1D) using Spearman's rank correlation, quantifying the model's internal representational alignment with theoretical predictions. To test whether dyadic effects were genuine, we generated 400 pseudo pairs by interleaving embeddings from different dyads while preserving trial order. This procedure disrupted dyad-specific interaction effects while maintaining intra-individual trial-level dynamics.

Representational complexity and misalignment

While RSA captures trial-by-trial similarity, intrinsic dimensionality (ID) provides a complementary measure of the number of dimensions a dyad utilizes to encode interaction histories. To quantify this metric of representational complexity, the participation ratio was employed (Gao et al., 2017):

$$ID = \frac{(\sum_i \lambda_i)^2}{\sum_i \lambda_i^2}$$

where λ represents the eigenvalues of the covariance matrix of the embeddings. A higher ID reflects a more distributed representation, whereas a lower ID indicates a more constrained representational subspace. To examine how interlocutors adapt their representations over time, differences in ID between dyadic partners were computed over expanding windows of trials:

$$\Delta ID = |ID_{\text{dyad},t} - ID_{\text{dyad},t-1}|$$

where $ID_{dyad,}$ represents the dyad's degree of representational misalignment at trial *t*. This metric captures the joint dynamics of representational complexity, providing a quantitative measure of the dyad's ability to align and stabilize representational frameworks across communicative exchanges.

Results

Task performance

Dyads successfully coordinated to arrange geometric shapes into target configurations, performing well above chance (~3%; eight locations with four potential orientations) across both datasets (TCG-SYNERGY1: NT = 80±9%, ASC = 80±11%; TCG-SYNERGY2: NT = 88±7%, ASC = 82±12%). While ASC and NT dyads performed comparably in TCG-SYNERGY1 (Welch's $t_{(49.5)} = 0.05$, $p > .05$, , $d_{Cohen} = 0.01$, 95% CI[-0.05, 0.05]), ASC dyads exhibited significantly lower performance in TCG-SYNERGY2 ($t_{(39.95)} = 2.2$, $p < .05$, $d_{Cohen} = 0.62$, 95% CI [0.01, 0.12]). Despite these differences, no significant group differences were observed in general task behaviors across both datasets, including planning time, movement time, or overall number of moves (all $t \leq .59$). Additionally, NT and ASC groups were matched in age, verbal IQ, and nonverbal IQ (all $t \leq .46$) but differed in AQ scores ($t_{(46.64)} = 12.05$, $p < .001$, $d_{Cohen} = 3.24$, 95%CI [14.38, 20.15]). These results confirm that the TCG is well within reach for both groups, as evidenced by their equivalent performance in TCG-SYNERGY1, where communicative problems followed a structured sequence. However, ASC dyads faced greater difficulty in TCG

SYNERGY2, where the unpredictable order of communicative problems placed greater demands on referential coordination, underscoring the task's sensitivity to ASC-related communicative challenges.

Interaction-level, not movement-level, embeddings differentiate NT from ASC dyads

The HTM, comprising movement- and interaction-level transformers along with a dyad-level classification head, demonstrated progressively improved classification performance across levels. At the movement level, the model classified dyads as NT or ASC at chance level (53% accuracy; binary precision: 51%; recall: 51%; F1-score: 51%), reinforcing previous findings showing matched movement-level behavior between NT and ASC dyads in the TCG (Wadge et al., 2019). At the interaction level, incorporating cumulative interaction history led to a significant improvement in classification accuracy (81% accuracy; binary precision: 80%; recall: 81%; F1-score: 80%). At the dyad level, pooling interaction-level embeddings further enhanced classification to 83% (binary precision: 84%; recall: 81%; F1-score: 82%). These improvements indicate that long-range dependencies in communicative behavior encode group differences more effectively than movement-level features, highlighting the importance of interaction-level dynamics in distinguishing NT and ASC dyads.

Interaction-level embeddings capture higher-order communication dynamics

To examine how the model represents communicative interactions, we projected movement- and interaction-level embeddings into a three-dimensional space using t-SNE with cosine distance (Figure 2B). At the movement level, embeddings were primarily organized by previously identified signal types (Wadge et al., 2019), with no clear separation between NT and ASC dyads, consistent with the model's chance-level classification performance at this stage. In contrast, interaction-level embeddings formed distinct NT and ASC clusters while preserving within-dyad signal organization, demonstrating the model's ability to encode higher-order communication dynamics essential for group differentiation.

Interaction-level embeddings encode communicative task structure

We further examined these group distinctions by correlating empirical dyad-specific interaction-level RSMs with hypothesis-driven similarity matrices capturing task-relevant referential organization, specifically *communicative problems* and *shape similarity* (Figure 1D). NT dyads exhibited stronger correlations with task constraints than ASC (Shape Similarity: $t_{(57.53)}= 5.58$, $p< .05$; Communicative Problems: $t_{(56.42)}=2.66$, $p<0.05$), with both groups differing from pseudo pairs (NT: $t_{(40.08)}=19.18$, $p<.001$ and $t_{(46.55)}=26.96$, $p<.001$; ASC: $t_{(30.95)}=12.37$, $p<.001$ and $t_{(34.70)}=22.93$, $p<.001$ for Shape Similarity and Communicative Problems, respectively; Figure 2C). Three additional analyses reinforce the specificity of this finding: (1) The observed effects reflect genuine dyadic interactions rather than statistical artifacts, as pseudo pairs exhibited near-zero correlations (Figure 2C). (2) The differences stem from cumulative interaction history, evident in the long-range dependencies in NT dyads' RSMs, which were less pronounced in ASD and pseudo pairs (Figure 2D). (3) The effect was exclusive to interaction-level embeddings, as movement-level embeddings failed to differentiate groups, reinforcing the observation that movement features primarily capture short-range dynamics without encoding broader communicative history.

Dynamic alignment of representational complexity within dyads shapes referential coordination

To understand how interlocutors regulate higher-order communication dynamics, we quantified representational misalignment as the absolute difference in ID between dyadic partners. ID estimates the number of dimensions utilized to encode interaction histories, providing insight into how interlocutors adapt their shared representational framework over time. In TCG-SYNERGY1, where communicative problems followed a structured sequence (Figure 1C), both NT and ASC dyads exhibited comparable representational alignment, maintaining low ID differences across trials. This suggests that when the problem order was predictable, both groups were able to stabilize their shared representational framework effectively. By contrast, pseudo pairs—randomly paired participants—displayed consistently higher misalignment, reflecting the absence of referential coordination in the absence of genuine interaction (Figure 3A). However, in TCG-SYNERGY2, where communicative problems were less predictable and introduced greater referential uncertainty, ASC dyads exhibited larger ID differences across trials than NT dyads. Notably, ASC misalignment overlapped with that of pseudo pairs for the first two-thirds of the task, indicating difficulties in adapting representational spaces to evolving communicative demands. In contrast, NT dyads stabilized referential alignment through reciprocal adaptation, even under volatile communicative conditions. These findings suggest that dynamic alignment of representational complexity is a control parameter supporting referential coordination.

Dynamic alignment of representational complexity predicts communicative success under referential volatility

The degree of representational misalignment was significantly associated with communicative performance. When dyads were categorized into low- and high-misalignment groups based on their ID differences in the final trial, NT dyads in the low-misalignment subgroup achieved a success rate of 87%, matching ASC dyads in the same subgroup (87%). However, ASC dyads in the high-misalignment subgroup performed significantly worse (76%) than NT dyads in the same category (90%; $t_{14.49} = 3.53$, $p < .01$. $d_{Cohen} = 1.4$, 95% CI [0.06, 0.23]; Figure 3D). This finding suggests that effective coordination of representational complexity within a dyad is crucial for adapting to unpredictable communicative dynamics. By contrast, in TCG-SYNERGY1, where problem order was structured, success rates did not significantly differ between high- and low-misalignment subgroups in either NT or ASC dyads (all $p > .65$; Figure 3C). This lack of differentiation suggests that the structured task design helped stabilize performance, mitigating the effects of misalignment on communicative success.

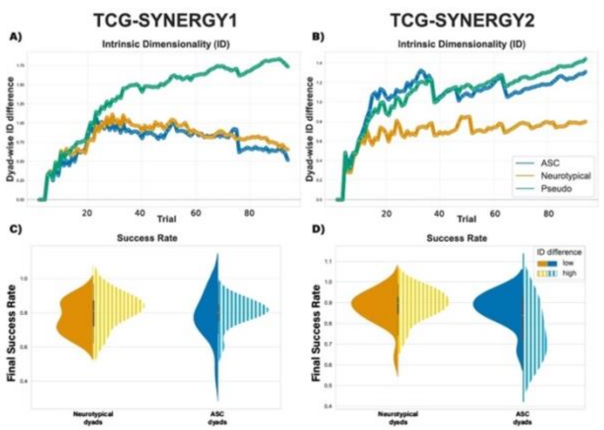

Figure 3: ID alignment dynamics and communicative success. **(A-B)** NT dyads (orange) and ASC dyads (blue) show comparable ID alignment in structured interactions (TCG-SYNERGY1), whereas ASC dyads exhibit persistent misalignment under referential volatility (TCG-SYNERGY2), mirroring pseudo pairs (green). **(C-D)** In TCG-SYNERGY1, communicative success is unaffected by ID misalignment, indicating stabilized referential alignment in structured sequences. In TCG-SYNERGY2, ASC dyads with higher ID misalignment perform significantly worse than NT dyads and low-misalignment ASC dyads.

Discussion

This study investigated how interlocutors coordinate novel signal-referent mappings across communicative turns in an experimental semiotic task (TCG), leveraging the known sensitivity of ASC dyads to referential volatility (Curcio & Paccia, 1987; Paul et al., 2009) to identify latent control parameters governing referential coordination. Using a custom HTM, we extracted multi-level embeddings, distinguishing movement-level features (within trials) from interaction-level features (across trials). Movement-level features failed to distinguish ASC from NT dyads (53% accuracy), confirming that individual behaviors were matched across groups. In contrast, interaction-level features differentiated ASC from NT dyads with 83% accuracy. Representational similarity analysis further revealed that interaction-level embeddings captured referentially relevant task dimensions through long-range dependencies, including across-trial problem structure. Moreover, interaction-level embeddings from NT dyads exhibited stronger alignment with task constraints than those of ASC dyads (16% vs. 9% of interaction-level variance). A key contribution of this study is the identification of representational complexity as a latent control parameter that may dynamically regulate referential coordination in context-dependent communication. Specifically, alignment in interaction-level representational complexity tracked communicative demands imposed by the semiotic challenge, with ASC dyads showing persistent misalignment under referential volatility. These findings provide empirical evidence for the notion that referential alignment is an interaction-level process that goes beyond statistical predictions of tokens, elucidating how interlocutors dynamically regulate referential complexity to meet communicative demands.

It has been argued that differences in basic kinematic properties of movement in individuals with ASC (Cook et al., 2009; Edey et al., 2016) may influence social interaction efficacy, potentially due to heterogeneity in motoric synchrony between NT and ASC individuals (Milton et al. 2022). However, our findings align with recent criticisms of the Interactional Heterogeneity Hypothesis (Georgescu et al., 2020; Mitchell et al., 2021), showing that communicative difficulties in linguistically-fluent ASC dyads arise from interaction-level processes rather than motoric variations. For instance, even when communicative performance and movement-level embeddings were matched between groups, interaction-level embeddings from NT individuals proved more sensitive to the referential structure of the semiotic challenge than those of ASC individuals. This suggests a disconnection between low- and high-level coordination features in communicative interactions, challenging accounts of communication that emphasize mutual priming and prediction across levels of linguistic organization (Garrod & Pickering, 2008; Pickering & Garrod, 2004).

It could be argued that the current findings cannot be generalized to linguistic communication, since the experimental constraints limit rapid multimodal interactions and repair procedures typical of everyday exchanges. In fact, these constraints were intentionally designed to amplify the semiotic challenges that human interlocutors routinely resolve when overcoming inter-individual differences in experience (Lupyan et al., 2023), conceptual representations (Marti et al., 2023), and prevalent signal-referent associations (Dor, 2015).

In summary, these findings suggest a computational mechanism implementing the notion of “common ground” (Clark, 1996), i.e. a continuously negotiated resource supporting mutual understanding in human communication. Specifically, we qualify how this resource is stabilized through referential alignment, achieved by matching the intrinsic representational dimensionality of a dyad’s interaction histories.

References

Christiansen, M. H., & Chater, N. (2022). The language game: How improvisation created language and changed the world. Random House.

Clark, H. H. (1996). Using Language. Cambridge University Press.

Contreras Kallens, P., Kristensen-McLachlan, R. D., & Christiansen, M. H. (2023). Large Language Models Demonstrate the Potential of Statistical Learning in Language. Cognitive Science, 47(3), e13256.

Cook, J., Saygin, A. P., Swain, R., & Blakemore, S.-J. (2009). Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. Neuropsychologia, 47(14), 3275–3278.

Curcio, F., & Paccia, J. (1987). Conversations with autistic children: contingent relationships between features of adult input and children's response adequacy. Journal of Autism and Developmental Disorders, 17(1), 81–93.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In arXiv [cs.CL]. arXiv. http://arxiv.org/abs/1810.04805

Dor, D. (2015). The instruction of imagination. Oxford University Press.

Edey, R., Cook, J., Brewer, R., Johnson, M. H., Bird, G., & Press, C. (2016). Interaction takes two: Typical adults exhibit mind-blindness towards those with autism spectrum disorder. Journal of Abnormal Psychology, 125(7), 879–885.

Galantucci, B., & Garrod, S. (2011). Experimental semiotics: a review. Frontiers in Human Neuroscience, 5, 11.

Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K., & Ganguli, S. (2017). A theory of multineuronal dimensionality, dynamics and measurement. In bioRxiv. bioRxiv. https://doi.org/10.1101/214262

Gareth Gaskell, M., & Altmann, G. (2007). The Oxford Handbook of Psycholinguistics. Oxford University Press.

Garrod, S., & Pickering, M. J. (2008). Shared circuits in language and communication. The Behavioral and Brain Sciences, 31(1), 26–27.

Georgescu, A. L., Koeroglu, S., Hamilton, A. F. de C., Vogeley, K., Falter-Wagner, C. M., & Tschacher, W. (2020). Reduced nonverbal interpersonal synchrony in autism spectrum disorder independent of partner diagnosis: a motion energy study. Molecular Autism, 11(1), 11.

Jazayeri, M., & Ostojic, S. (2021). Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. In arXiv [q-bio.NC]. arXiv. http://arxiv.org/abs/2107.04084

Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., & Krishnan, D. (2020). Supervised Contrastive Learning. In arXiv [cs.LG]. https://doi.org/10.48550/ARXIV.2004.11362

Lepori, M. (2020). Unequal representations: Analyzing intersectional biases in word embeddings using representational similarity analysis. Proceedings of the 28th International Conference on Computational Linguistics. Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain (Online). https://doi.org/10.18653/v1/2020.coling-main.151

Levinson, S. C. (2020). On the human "interaction engine." In Roots of Human Sociality (pp. 39–69). Routledge.

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. Nature Reviews. Disease Primers, 6(1), 5.

Lupyan, G., Uchiyama, R., Thompson, B., & Casasanto, D. (2023). Hidden differences in phenomenal experience. Cognitive Science, 47(1), e13239.

Marti, L., Wu, S., Piantadosi, S. T., & Kidd, C. (2023). Latent diversity in human concepts. Open Mind, 7, 79–92.

Milton, D., Gurbuz, E., & Lopez, B. (2022). The "double empathy problem": Ten years on. Autism : The International Journal of Research and Practice, 26(8), 1901–1903.

Mitchell, P., Sheppard, E., & Cassidy, S. (2021). Autism and the double empathy problem: Implications for development and mental health. The British Journal of Developmental Psychology, 39(1), 1–18.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. In arXiv [cs.LG]. https://doi.org/10.48550/ARXIV.1912.01703

Paul, R., Orlovski, S. M., Marcinko, H. C., & Volkmar, F. (2009). Conversational behaviors in youth with high-functioning ASD and Asperger syndrome. Journal of Autism and Developmental Disorders, 39(1), 115–125.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. The Behavioral and Brain Sciences, 27(2), 169–190; discussion 190–226.

Reddy, M. J. (1993). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), Metaphor and Thought (pp. 164–201). Cambridge University Press.

Roads, B. D., & Love, B. C. (2024). The dimensions of dimensionality. Trends in Cognitive Sciences, 28(12), 1118–1131.

Stolk, A., Verhagen, L., Schoffelen, J.-M., Oostenveld, R., Blokpoel, M., Hagoort, P., van Rooij, I., & Toni, I. (2013). Neural mechanisms of communicative innovation. Proceedings of the National Academy of Sciences of the United States of America, 110(36), 14574–14579.

Stolk, A., Verhagen, L., & Toni, I. (2016). Conceptual Alignment: How Brains Achieve Mutual Understanding. Trends in Cognitive Sciences, 20(3), 180–191.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In arXiv [cs.CL]. arXiv. http://arxiv.org/abs/1706.03762

Wadge, H., Brewer, R., Bird, G., Toni, I., & Stolk, A. (2019). Communicative misalignment in Autism Spectrum Disorder.

Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 115, 15–26.