Fine-tuning conceptual structure of referents through coordinated interaction

Miriam Poncelet-Romaneli (miriam.greidanusromaneli@donders.ru.nl)
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

Thibaud Glasman (thibaud.glasman@ru.nl)
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

Amir Homayun Hallajian (amirhomayun.hallajian@donders.ru.nl)
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

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

Bruno Galantucci (bruno.galantucci@gmail.com)
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

Abstract

During communicative interactions, language production and comprehension are bounded by the accumulation of a shared communicative context. Besides lexical pacts and simplification of referential expressions, the creation of a shared context allows for the intelligible use of linguistic signals with novel, interaction-specific meanings. Here, we explore whether context-specific language use leads to mutual adjustment of interlocutors' conceptual representations. We tasked dyads with solving a referential communication game, quantifying dialogue-related adjustments in interlocutors' conceptual structures, and coordination dynamics during the interaction. After engaging in the dialogue, interlocutors judged the same set of referents more similarly than other participants engaged in the same task, but not with each other (pseudo-pairs). Exploratory analyses of the structural complexity of unfolding semantic spaces indicate a stronger alignment between interacting dyads than pseudo-pairs. These findings suggest that human communication is supported by structural coordination of conceptual representations of the communicative referents, over and above signal-level alignment.

Keywords: referential communication; conceptual spaces; coordination; intrinsic dimensionality; Tangram task

Introduction

Flexibility is a pervasive feature of how linguistic signals (e.g., words, utterances) map onto the mental states they evoke (Casasanto & Lupyan, 2015; Elman, 2004; Ludlow, 2014; Lupyan & Lewis, 2017). We adapt how we produce and interpret signals based on our communicative goals (Grice, 1957), our interlocutors' social group memberships (Clark & Murphy, 1982), as well as on the accumulating context emerging over the course of an interaction. Previous work on referential communication has shown that such an interaction-specific shared context allows interlocutors to produce increasingly underspecified signals (Clark & Wilkes-Gibbs, 1986), which interlocutors, but not overhearers, can nonetheless interpret (Schober & Clark, 1989). These observations suggest that resolving underspecified reference requires more than introducing and reusing signals. First, interlocutors surmount short-range communicative obstacles: at each turn, they need to detect and jointly overcome individual differences in experience (Lupyan, Uchiyama, Thompson, & Casasanto, 2023), conceptual representations (Marti, Wu, Piantadosi, & Kidd, 2023), and signal-referent associations (Dor, 2015). Second, interlocutors also need to surmount long-range obstacles: across distant turns, interlocutors generate a joint interpretational framework to disambiguate the evolving signal-referent associations they use (Stolk, Bašnáková, & Toni, 2020).

The coordination between interlocutors to build a communicative context has been extensively described with the help of referential communication games, including variants of the Tangram task (Clark & Wilkes-Gibbs, 1986; Krauss & Weinheimer, 1964). However, the focus of investigation has predominantly concentrated on effects of communicative interactions on signals themselves, such as the creation of lexical pacts, the simplification of referential expressions, and the persistence of contextually disambiguating signals (Brennan & Clark, 1996; Hawkins, Goodman, Goldberg, & Griffiths, 2020). Less attention has been devoted to measuring the mutual adjustment of interlocutors' conceptual representations that allows for signals to acquire novel, interaction-specific meanings. Here, we build on theoretical suggestions that linguistic signals function by discretizing continuous experience (Dor, 2015), and that good-enough alignment in the structure of semantic spaces can support communication (Gärdenfors, 2014; Warglien & Gärdenfors, 2015). Accordingly, rather than focusing on the signals people use to represent a set of

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

communicative referents, we focus on the conceptual representations that support the creation and use of such signals. By investigating structural conceptual adjustments, we aim to gain understanding of how interlocutors use signals to concurrently converge on their use, and on the underlying discretization of their experience of the world. This can be contrasted with the more conventional approach of ignoring these situated conceptual representations and focus on *a priori* features of conceptual spaces (Binder et al., 2016; McRae, de Sa, & Seidenberg, 1997). This approach has offered valuable insights, but might not be well-tuned to capture the idiosyncratic signal-referent relations that emerge through interaction, as the relevant features for representation may vary across individuals (Marti et al., 2023) and are themselves conveyed through language.

Referential communication games have already shown to create and foster similarities between interlocutors, making dyads generate signals which, over repeated use, not only become more similar lexically and syntactically (Branigan, Pickering, & Cleland, 2000; Eijk et al., 2022), but that can also be used to refer to a set of novel stimuli (Eliav, Ji, Artzi, & Hawkins, 2023). Transfer might indicate that signal alignment is a manifestation of an underlying conceptual alignment extending beyond the specific referents used during the Tangram task. However, transfer does not directly capture mutual adjustments in interactants' conceptual representations of the communicative referents.

Here, we operationalize dialogue-related conceptual adjustments by considering the *structure* of interlocutors' conceptual representations of a set of Tangram stimuli, rather than arbitrarily selecting diagnostic features of those representations. To this end, we measure similarity and ambiguity relationships between referents in a closed set of stimuli, tracking how those relationships are shaped by communicative interactions between pairs of participants. Furthermore, we explore whether the dialogue dynamics influences these conceptual adjustments between interlocutors. Namely, we explore whether inter-dyad variation in communicative effort (Bögels et al., 2024), semantic and acoustic coordination (Abney, Paxton, Dale, & Kello, 2021; O'Nell & Finn, 2024; Van Orden, Kloos, & Wallot, 2011), as well as alignment in representational structure (Jazayeri & Ostojic, 2021) accounts for the degree of dyads' mutual conceptual adjustments.

Methods

This report is based on exploratory analyses designed to generate hypotheses and effect size estimates to be subsequently evaluated in a confirmatory follow-up study. Thirty-six participants were paired at random into 18 dyads, after providing informed consent (Ethics Committee Social Sciences, Radboud University), with one dyad excluded before data acquisition, and three dyads lacking part of the data., resulting in a final sample of 14 dyads The two participants in a dyad were seated in separate rooms, linked only by an audio channel. Participants completed a *referential communication task* and a *joint labelling task*

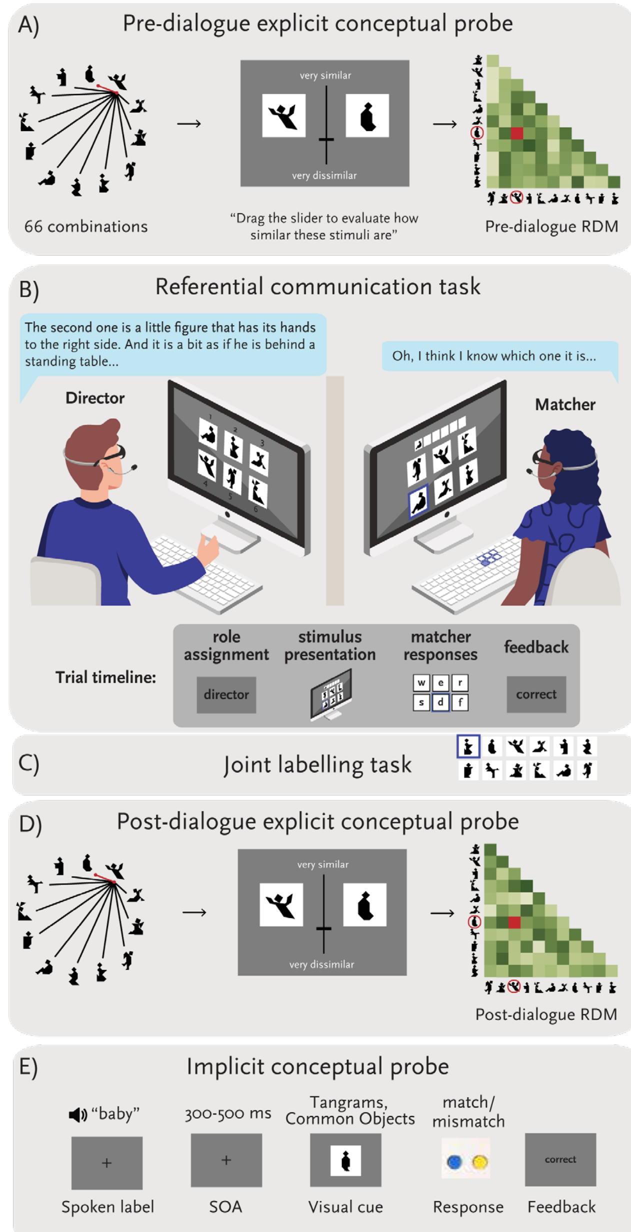

Figure 1. Experimental procedure. A) *Pre-dialogue explicit conceptual probe*: Before engaging in the referential communication task, each participant provided (individually) a similarity judgement over each combination of 12 Tangram stimuli, resulting in a participant-specific pre-dialogue representational dissimilarity matrix (RDM). B) *Referential communication task:* On each of 24 trials, members of a dyad alternated in the Director/Matcher roles, coordinating in sorting a set of 6 stimuli in the order shown to the Director. C) *Joint labelling task:* Dyads generated a single word label for each of the 12 Tangrams. D) *Post-dialogue explicit conceptual probe:* This task, identical to A), provided a post-dialogue RDM of the Tangrams. E) *Implicit conceptual probe:* Individual participants reported matches/mismatches between spoken labels and visual stimuli involving Tangrams and Common Objects.

(jointly, audio channel on), as well as two conceptual probe tasks (individually, audio channel off): an *explicit conceptual probe,* performed before and after the referential communication task, and an *implicit conceptual probe*, performed only after the referential communication task. Figure 1 shows the tasks as completed in chronological order.

Tasks

Referential Communication Task. During each trial of the referential communication task (Figure 1B), participants saw a set of 6 Tangram stimuli sampled from a set of 12 Tangrams (corresponding to those used by Clark & Wilkes-Gibbs, 1986), displayed on-screen in a 2x3 grid. Each of the participants was assigned the role of director or matcher, which alternated every other trial. Director and matcher were allowed to communicate freely throughout the task through an audio channel, which was recorded, but were not able to see each other during the interaction. The director was given private information about the target order of the stimuli, corresponding to the positions of the stimuli on the grid (from top-left to bottom-right), and marked by numbers 1-6. The joint goal of the task was for the matcher, who saw the stimuli in randomly scrambled positions, to select the stimuli in the order corresponding to that seen by the director. The matcher could select stimuli by making a keyboard press on an array of keys corresponding to the on-screen stimulus positions. Stimuli already selected by the matcher were displayed to them (but not to the director) on a smaller 1x6 array above the main stimulus grid, as well as marked by a blue frame around the selected stimulus on the main grid. Matchers were allowed a mechanism for response "repair" by pressing the spacebar, which erased all previous selections made during a trial. Trials automatically ended once the matcher selected all six on-screen stimuli, after which both participants were simultaneously displayed with feedback for their performance: "correct" in case the selection order corresponded to the target order, and otherwise "incorrect". The task consisted of 24 trials, matched for frequency of occurrence of the 12 stimuli, as well as for the frequency of each stimulus among the first three stimuli in a trial.

Joint labelling task. Only after the referential communication task was completed, dyads were informed that they should jointly assign a single-word descriptive label to identify each of the stimuli. During this labeling phase, the full set of stimuli was presented at once to the participants and, one at a time, the stimulus to be labeled was highlighted (Figure 1C). Dyads could freely discuss how to label the stimuli, alternating which participant was tasked with typing in the label. For each dyad, the typed labels for each stimulus were immediately converted into spoken labels using a text-to-speech model (eleven_multingual_v2). Dyad-specific spoken labels were used as auditory cues during the implicit conceptual probe. Labels corresponded predominantly (91%) to expressions used by the dyads during dialogue.

Explicit Conceptual Probe. Before (Figure 1A) and after (Figure 1D) the referential communication task, participants were asked to individually provide explicit similarity judgements over the 12 Tangram stimuli. We presented participants with each of the 66 unique pairwise stimulus combinations in this set, visually displaying the two stimuli on screen and asking participants to drag a slider to evaluate "how similar or dissimilar these stimuli are". The end points of the sliding scale were labeled "very similar" and "very dissimilar", and each trial began with the slider on the dissimilarity extreme of the scale. Trials were self-paced, and the order of presentation of the stimulus combinations was randomized across participants. This identical procedure was conducted both pre- and post-dialogue.

Implicit Conceptual Probe. After the referential communication task, the joint labelling task, and the explicit conceptual probe, participants were asked to individually report matches or mismatches between pair-specific labels for the Tangram stimuli, presented as auditory cues, and the corresponding visual Tangrams (Figure 1E). To evaluate whether within-dyad similarities in performance during this task were specifically linked to referents included in their previous dialogue (i.e. the Tangram stimuli), we asked participants to perform the same task using highly-nameable drawings of Common Objects that were not included in the dialogue, but were likely highly familiar to the participants. The drawings were extracted from a normed dataset (Snodgrass & Vanderwart, 1980), and selected to match representative semantic domains of the noun phrases produced at the end of a similar Tangram task (Hawkins, Frank, & Goodman, 2020). On each trial of the implicit conceptual probe, participants listened to one spoken word over headphones, and reported a match/mismatch with the visual stimulus presented on a computer monitor after a randomized lag of 300-500 ms following the end of the spoken word, by pressing their left/right index fingers on a custom button-box. Match/mismatch mappings with left/right index fingers were counterbalanced across participants. The task consisted of 504 base trials, 240 match trials and 264 mismatch trials, arranged in four blocks counterbalanced between Tangrams and Common Objects, and between participants. Incorrect trials from the first two blocks were repeated in the final two blocks, with a total number of completed match trials varying between 240 and 286 trials across participants.

Data analysis

Referential Communication Task. To quantify coordination within a dyad during the dialogue, we considered three parameters: communicative effort, alignment in representational structure, and coordination dynamics (see below). Audio recordings of each dyad's dialogue were recorded as stereo files, with recordings from each participant on separate audio tracks. Stereo tracks were split, transcribed with timestamps using WhisperX large-v3 model, and recombined into a time-stamped dialogue transcription for each dyad. The transcribed dialogue was checked and amended by a native Dutch speaker, and segmented into communicative turns, each turn consisting of consecutive language production by a given participant. Each communicative turn was then embedded in the high-dimensional space of a Sentence Transformer model (all-mpnet-base-v2).

Communicative effort. Following the well-established pattern of reduction in referential expressions during the Tangram Task (Clark & Wilkes-Gibbs, 1986), we estimated how efficiently interacting dyads converge on shortened referential expressions (Bögels et al. 2024). More precisely, we fit a negative power function to the trial-wise word count produced by a dyad over the dialogue, according to the function $y = a \cdot x^{-c}$, with y as the number of words produced, and x as the trial number. The resulting parameters c and a reflect, respectively, the steepness of the decline in communicative effort over the course of the dialogue, and the overall communicative effort exerted by the interlocutors.

Representational structure alignment. Inspired by suggestions that communication is supported by similarity between the structures of interlocutors' semantic spaces (Gärdenfors, 2014), we computed the intrinsic dimensionality (ID), a measure of representational complexity (Jazayeri & Ostojic, 2021; Roads & Love, 2024), of the cumulative semantic spaces built from the semantic embeddings of individual dialogue turns. The ID of a semantic space corresponds to the participation ratio of the eigenvalues (λ) of the representational similarity matrix of its semantic embeddings, which is constructed by computing the cosine similarity between dimensions of the embeddings. (Sorscher, Ganguli, & Sompolinsky, 2022):

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

To quantify (mis)alignment in representational structure between interlocutors over the course of the interaction, we constructed an ID trajectory for each interlocutor based on the ID values of semantic spaces containing expanding windows of turns produced by this interlocutor. Note that ID can decrease when subsequent turns are redundant or constrain the spectrum of λ. By computing the sum of the distances between the ID trajectories of two interlocutors in a dyad (ΔID) over each adjacent pair of turns t, we aimed to capture fluctuations in representational complexity between interlocutors, dynamically measuring their ability to establish and maintain a cohesive shared framework across successive communicative exchanges.

$$\Delta ID = \sum_t \left| ID_{Speaker\ A,t} - ID_{Speaker\ B,t} \right|$$

Coordination dynamics. Lastly, we quantified the magnitude of long-range dependencies and self-affinity in the dialogues of each dyad, to evaluate whether the dialogues evoked during this experimentally contrived situation reflect the coordination dynamics reported in dialogue under naturalistic conditions (Alviar, Kello, & Dale, 2023; O'Nell & Finn, 2024). After calculating the distribution of turn-by-turn distances in semantic embeddings generated by the Sentence Transformer model (*semantic synchrony*; O'Nell & Finn, 2024), the degree of long-range correlation of each dyad's semantic synchrony timeseries was assessed using detrended fluctuation analysis (DFA; Rydin Gorjão, Hassan, Kurths, & Witthaut, 2022). DFA allows for the analysis of non-stationary timeseries by accounting for signal trends at varying scales. Here, we measured variance in dyads' semantic synchrony timeseries at different time scales. The scaling coefficient (α) of a timeseries was estimated by fitting a power function $F(s) \propto s^{\alpha}$, where *F(s)* is the fluctuation function and s is the length of the windows used to calculate those fluctuations. This scaling coefficient characterizes the nature of the long-range correlations, with $\alpha = 0.5$ indicating uncorrelated data (random fluctuations, i.e. white noise) and $\alpha > 0.5$ indicating increasingly long-range correlations (Kantelhardt, Koscielny-Bunde, Rego, Havlin, & Bunde, 2001).

To validate the robustness of the scaling coefficient obtained from the dialogues, the observed coefficient was compared to a null distribution of scaling coefficients, generated through scrambling the signal's temporal order. This procedure results in a Pink Noise Robustness (PNR) value, the proportion of the null distribution below the observed coefficient (O'Nell & Finn, 2024), providing a signature of adaptive coordination between speakers. By comparing the observed scaling exponent and shuffled baselines with disrupted temporal dependencies, PNR scores help dissociate *i)* the self-affinity in the semantic synchrony signal attributable to mutual adaptations between speakers from *ii)* confounds such as self-affinity resulting from extended sections of low synchrony signal.

Explicit similarity probe. This behavioral task was designed to test the prediction that before the dialogue, real dyads and pseudo-pairs would have matched similarity judgments, whereas after the dialogue, similarity judgements in real dyads would be closer than those in pseudo-pairs. For each participant and session (pre- versus post-dialogue), we constructed a representational dissimilarity matrix (RDM) between stimuli based on the explicit similarity ratings given by the participant for each combination of two stimuli. We then computed the Spearman (rank) correlation between two given matrices as a measure of similarity between participants' representations of the dissimilarity between the stimuli. To test whether the correlation of dyads' explicit judgements increased above and beyond the potential effect of familiarization with the stimuli across the sample, we compared the similarity of dyads with that of pseudo-pairs, pairwise combinations of participants in the sample who did not interact with each other. Given the asymmetric group sizes between real dyads (*n*=17) and pseudo-pairs (*n*=378), we tested differences in pairwise similarity between groups using a non-parametric permutation analysis, where we randomly assigned participant combinations to one of the two groups, and created a distribution of t-statistics by comparing the means of the shuffled groups over 100,000 permutations (Chen et al., 2016). The resulting permutation distributions provided us with a confidence interval over which we could then infer differences between groups.

Implicit similarity probe. This behavioral task was designed to test the prediction that correlations in signal-referent ambiguities would be higher in real pairs than in pseudo-pairs for the Tangrams, but not for familiar stimuli not included in the dialogue (Common Objects). Reaction times were calculated as the time interval between the presentation of the visual cue and the onset of the button press. Reaction times shorter than 200 ms or longer than 1200 ms were removed as outliers. For each participant, native reaction times were residualized against a number of confound variables that could have contributed to the participants' responses. More precisely, for each stimulus, we considered the following confound variables: length of the spoken word (*audio length*); pixel-wise similarity of a visual stimulus against the rest of the set (separately for Tangrams and Common Objects; *visual ambiguity*); cosine similarity between semantic embeddings of a stimulus label against the rest of the set (using all-mpnet-base-v2; *semantic ambiguity*); Levenshtein distance between labels and the rest of a dyad's label set (*word-form ambiguity*). We account for potential effects of these confounds variables on reaction times by fitting a linear regression model of log-transformed reaction times with confound variables as predictors, and participants as random effects (log reaction times ~ audio length + visual ambiguity + semantic ambiguity + word-form ambiguity + (1|participant)). Subsequent analyses used the model residuals, and focused on the matched trials, where we expected that reaction times to a given signal-referent mapping increased proportionally to the ambiguity of this signal-referent pairing relative to the other signals and referents.

Results

Communicative coordination: effort dynamics. Word production decreased across trials of the referential communication task following a negative power law (overall effort: a=441, range = [260, 751]; decreasing effort speed: c=0.75, range = [0.60, 1.04]; R^2=0.98; Figure 2A). This measure provides a dyad-specific parametrization of communicative effort, confirming that it follows a power-law dynamics (Bögels et al., 2024).

Communicative coordination: dynamic representational alignment. There was an initial increase in ID_t from the semantic embeddings generated by each interlocutor during the dialogues. Some dyads had closely matched ID_t profiles (Figure 2B, left panel), whereas other dyads diverged, particularly so in the second half of the dialogue (Figure 2B, right panel). Overall, trials' final ΔID in real dyads (11.72±7.77, mean±sd) was significantly lower than ΔID in pseudo-pairs (20.56±13.64, z=-2.71, p=.007; Figure 2C). We accounted for the possible confounding effect of word production by running a linear regression with group type (real dyads/pseudo pair) and word production as predictors. The group type was a significant predictor of ΔID (p=.007).

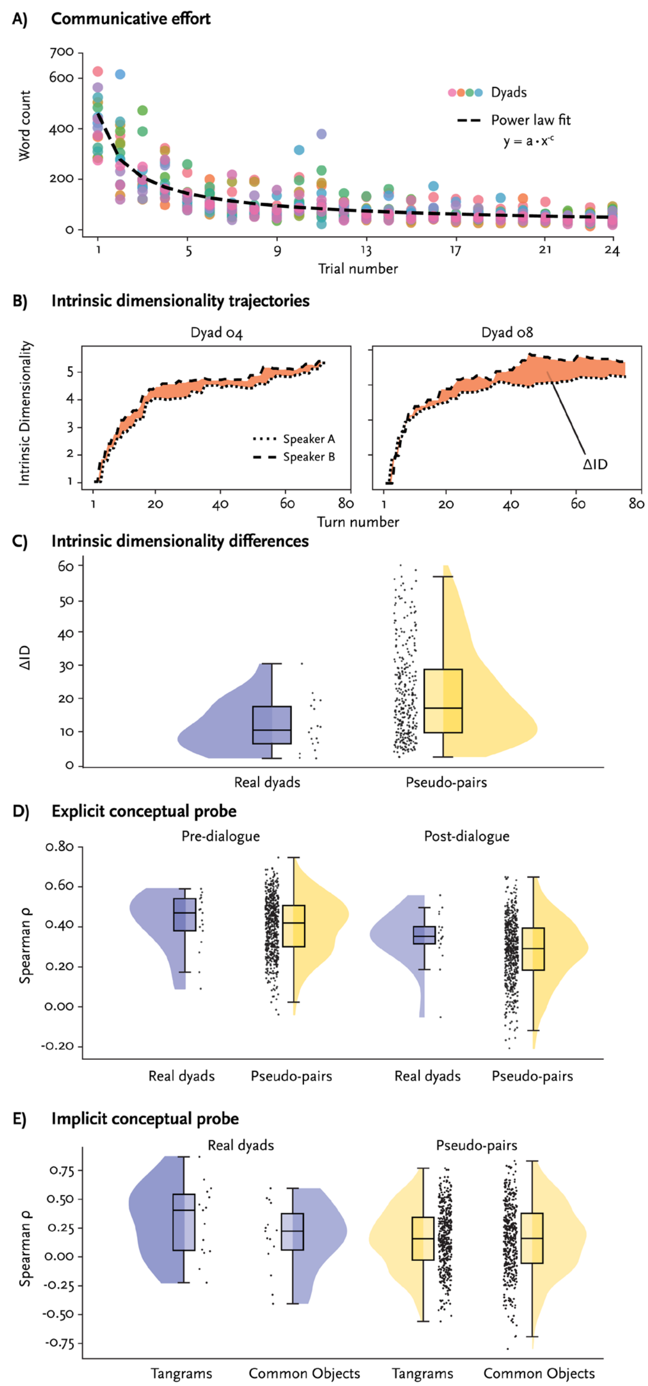

Figure 2. A) *Communicative effort dynamics*. The number of words produced per trial decreases according to a power function. B) *Intrinsic Dimensionality trajectories.* Examples of within-dyad ID_t. C) *Intrinsic dimensionality differences.* ΔID is lower in real dyads than in pseudo-pairs. D) *Explicit conceptual probe.* After the dialogue, but not before, dyads evaluate Tangrams more similarly than pseudo-pairs. E) *Implicit conceptual probe.* Real dyads take similarly long to match spoken labels and visual referents for Tangrams they communicated about, but not for drawings of familiar Common Objects that were not part of their dialogue.

This indicates that, above and beyond the variety of dyad-specific ID_t dynamics, real dyads coordinate their mutual distance in ID trajectories over time.

Conceptual adjustments: explicit probe. Before the dialogues, explicit similarity judgements are similarly correlated in real dyads (0.43±0.14) and in pseudo-pairs (0.40±0.14; z=0.68, p=.24; Figure 2C). After the dialogues, explicit similarity judgements become more similar in real dyads (0.35±0.14) than in pseudo-pairs (0.28±0.16; z=1.61, p=.04). This finding indicates that, after the dialogues evoked by the Tangram task, real dyads think of the referents in a manner more similar than pseudo-pairs.

Conceptual adjustments: implicit probe. Reaction times to the Tangrams' labels are more correlated in real dyads (0.28±0.31) than in pseudo-pairs (0.10±0.27; z=2.49, p<.01; Figure 2D). Reaction times to the Common Objects are matched between real dyads (0.25±0.20) and pseudo-pairs (0.26±0.27; z=-0.06, p=.47). This finding indicates that pairs of participants take similarly long to match spoken labels to their visual referents, but only when participants jointly coordinated on those labels, not when the referents were not part of their dialogues nor when participants where not part of the same dyad.

Communicative coordination: pink noise robustness. The scaling coefficient of a DFA over semantic synchrony extracted from each dialogue (DFA_α: 0.65±0.11) was significantly higher than the average null distribution of temporally scrambled semantic synchrony timeseries (α_{null}=0.54; PNR: 0.79±0.22, i.e. 79% of scrambled coefficients is lower than DFA_α, p<.05). This finding indicates the presence of pink noise dynamics in the semantic synchrony of the real dyads, a signature of adaptive coordination between speakers.

Relations between communicative coordination and conceptual adjustments. Exploratory analyses indicate the possibility of a relation between the amount of effort exerted by a dyad during the dialogue and how similarly members of a dyad think about Tangrams relations after the dialogue. The intercepts of the power function fitted to the pattern of reduction in word count of each dialogue are negatively associated with dyads' similarity in explicit probe responses after the dialogue ($\rho_{(15)}$=-0.52, p=.03). To detect this effect at 80% statistical power and α=.05, a confirmatory sample would need 18 dyads. A second exploratory analysis indicates the possibility of a relation between the ID distance developed within a dyad during the dialogue, and how similarly members of a dyad think about Tangrams relations after the dialogue. The sums of ΔID over turns of each dialogue are negatively associated with dyads' similarity in explicit probe responses after the dialogue ($\rho_{(15)}$=-0.36, p=.16). To detect this effect at 80% statistical power and α=.05, a confirmatory sample would need 56 dyads. We accounted for participants' word production by running a linear model with word counts and ΔID as predictors of the similarity in probe responses. Considering a possible interaction between the predictors, a power analysis suggests a required sample size of 51 dyads to detect this effect. These findings suggest that metrics of within-dialogue coordination are consequential for how similarly members of a dyad come to think of the Tangrams.

Discussion

In this exploratory study, we proposed to look beyond known effects of dialogue on lexical alignment, to understand how interlocutors coordinate conceptual adjustments to communicative challenges. As dialogues progressed, we measured different aspects of the coordination dynamics between the interlocutors, from their communicative effort, to their coordination of semantic dynamics (PNR), to the complexity-matching between unfolding semantic spaces they produced (ΔID). We then tested whether and how the conceptual representations that support the creation and use of the signals in dialogue were modified by the referential coordination process evoked by the Tangram Task.

We provided empirical evidence for the notion that dialogue may lead to increased alignment between interlocutors' conceptual structures. Specifically, we found that, after the dialogue, real dyads judged the same set of referents as more similar than pseudo-pairs (i.e. participants engaged in dialogue, but not with each other). Real dyads were also more similar than pseudo-pairs in responding to signal-referent mappings involving the referents of the dialogue, but not when responding to objects for which they shared no joint experiences.

Additionally, we identified a potential latent parameter supporting the referential coordination occurring in real dyads, namely, interlocutors' alignment in the intrinsic dimensionality used to encode their interaction histories. More precisely, we found within-dyad differences in the intrinsic dimensionality of the semantic embeddings supporting the dialogue of each dyad, but those differences were considerably smaller than those measured between pseudo-pairs. We also observed a possible relationship at the dyad level between this parameter and the extent of conceptual alignment between interlocutors, which we will further investigate in a confirmatory replication of this study.

Overall, the current findings open the possibility of understanding referential communication in terms of quantitative relations between specific dynamical metrics of communicative coordination and dyad-specific conceptual adjustments. The exploratory analyses of this study suggest that control over fluctuations in intrinsic dimensionality differences between interlocutors, as well as the dynamics of communicative effort, might be consequential for interlocutors' conceptual coordination. It remains to be seen whether these relations can be statistically replicated in a suitable confirmatory sample, as well as how they can be neurally implemented.

References

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2021). Cooperation in sound and motion: Complexity matching in collaborative interaction. *Journal of Experimental Psychology: General*, *150*, 1760–1771.

Alviar, C., Kello, C. T., & Dale, R. (2023). Multimodal coordination and pragmatic modes in conversation. *Language Sciences*, *97*, 101524.

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, *33*, 130–174.

Bögels, S., Li, T., Rasenberg, M., Eijk, L., Toni, I., & Pouw, W. (2024, June 14). *There is a power law of joint communicative effort and it reflects communicative work*. OSF.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*, B13–B25.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482.

Casasanto, D., & Lupyan, G. (2015). All Concepts Are Ad Hoc Concepts. In E. Margolis & S. Laurence (Eds.), *The Conceptual Mind* (pp. 543–566). The MIT Press.

Chen, G., Shin, Y.-W., Taylor, P. A., Glen, D. R., Reynolds, R. C., Israel, R. B., & Cox, R. W. (2016). Untangling the relatedness among correlations, part I: Nonparametric approaches to inter-subject correlation analysis at the group level. *NeuroImage*, *142*, 248–259.

Clark, H. H., & Murphy, G. L. (1982). Audience Design in Meaning and Reference. *Advances in Psychology*, *9*, 287–299.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Dor, D. (2015). *The Instruction of Imagination: Language as a Social Communication Technology*. Oxford University Press.

Eijk, L., Rasenberg, M., Arnese, F., Blokpoel, M., Dingemanse, M., Doeller, C. F., … Bögels, S. (2022). The CABB dataset: A multimodal corpus of communicative interactions for behavioural and neural analyses. *NeuroImage*, *264*, 119734.

Eliav, R., Ji, A., Artzi, Y., & Hawkins, R. D. (2023, May 10). *Semantic uncertainty guides the extension of conventions to new referents*. arXiv.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, *8*, 301–306.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press.

Grice, H. P. (1957). Meaning. *The Philosophical Review*, *66*, 377–388.

Hawkins, R. D., Frank, M. C., & Goodman, N. D. (2020). Characterizing the Dynamics of Learning in Repeated Reference Games. *Cognitive Science*, *44*, e12845.

Hawkins, R. D., Goodman, N. D., Goldberg, A. E., & Griffiths, T. L. (2020). Generalizing meanings from partners to populations: Hierarchical inference supports convention formation on networks. *arXiv:2002.01510 [Cs]*. Retrieved from http://arxiv.org/abs/2002.01510

Jazayeri, M., & Ostojic, S. (2021). Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *Current Opinion in Neurobiology*, *70*, 113–120.

Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S., & Bunde, A. (2001). Detecting long-range correlations with detrended fluctuation analysis. *Physica A: Statistical Mechanics and Its Applications*, *295*, 441–454.

Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. *Psychonomic Science*, *1*, 113–114.

Ludlow, P. (2014). *Living Words: Meaning Underdetermination and the Dynamic Lexicon*. Oxford University Press.

Lupyan, G., & Lewis, M. (2017, September 4). *From words-as-mappings to words-as-cues: The role of language in semantic knowledge*. PsyArXiv.

Lupyan, G., Uchiyama, R., Thompson, B., & Casasanto, D. (2023). Hidden Differences in Phenomenal Experience. *Cognitive Science*, *47*, e13239.

Marti, L., Wu, S., Piantadosi, S. T., & Kidd, C. (2023). Latent Diversity in Human Concepts. *Open Mind*, *7*, 79–92.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). *On the Nature and Scope of Featural Representations of Word Meaning*.

O'Nell, K. C., & Finn, E. S. (2024). Pink noise in speakers' semantic synchrony dynamics as a metric of conversation quality. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*. Retrieved from https://escholarship.org/uc/item/3q474016

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

Rydin Gorjão, L., Hassan, G., Kurths, J., & Witthaut, D. (2022). MFDFA: Efficient multifractal detrended fluctuation analysis in python. *Computer Physics Communications*, *273*, 108254.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*, 211–232.

Snodgrass, J., & Vanderwart, M. (1980). A standardized set of 260 pictures. Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *6*, 174–215.

Sorscher, B., Ganguli, S., & Sompolinsky, H. (2022). Neural representational geometry underlies few-shot concept learning. *Proceedings of the National Academy of Sciences*, *119*, e2200800119.

Stolk, A., Bašnáková, J., & Toni, I. (2020). Joint epistemic engineering: The neglected process of context

construction in human communication. In *Routledge Handbook of Neurosemiotics* (pp. 259–278). Routledge.
Van Orden, G. C., Kloos, H., & Wallot, S. (2011). Living in the Pink. In *Philosophy of Complex Systems* (pp. 629–672). Elsevier.
Warglien, M., & Gärdenfors, P. (2015). Meaning Negotiation. In F. Zenker & P. Gärdenfors (Eds.), *Applications of Conceptual Spaces* (pp. 79–94). Springer International Publishing.