Resolving Communicative Uncertainty through Computational Inference of Partner Intentions

Tengfei Zhang† (tengfei.zhang@mail.bnu.edu.cn)
State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research,
Beijing Normal University, Beijing 100875, China

Chujie Ye† (yechujie@mail.bnu.edu.cn)
State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research,
Beijing Normal University, Beijing 100875, China

Lusha Zhu* (lushazhu@pku.edu.cn)
School of Psychological and Cognitive Sciences; Beijing Key Laboratory of Behavior and Mental Health; IDG/McGovern Institute for Brain Research;
Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

Chunming Lu* (luchunming@bnu.edu.cn)
State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research,
Beijing Normal University, Beijing 100875, China

†These authors contributed equally to this work. *Corresponding authors.

Abstract

Effective social communication demands that individuals adeptly align their conceptual representations through the precise use of language. Yet, how individuals resolve uncertainty in selecting context-appropriate utterances remains a core question in cognitive science and a significant challenge for large language models (LLMs). In this study, 60 participants (30 same-gender dyads) performed a collaborative word generation task designed to capture the dynamics of open-ended, two-way communication. Our results show that human interlocutors can effectively resolve communicative uncertainty and achieve mutual understanding, even in unconstrained, ambiguous exchanges. Furthermore, drawing on established psycholinguistic theories, we developed computational models within the cohort-based, selection-by-competition framework to test two competing mechanisms. The findings suggest a functional division of labor: statistical learning (SL) facilitates the generation of candidate lexical cohorts, while pragmatic reasoning (PR) predominantly governs word selection within the cohort.

Keywords: collaborative communication; pragmatic reasoning; statistical learning; cognitive modeling; large language model

Background

Effective social communication requires individuals to align their conceptual representations through precise utterance selection. However, in everyday two-way verbal communication, the process of communication is often broken down due to the inherent ambiguity of utterances regarding the intended meanings of the partner. This ambiguity mainly arises from the influence of prior context and its dynamic update over time (Hawkins et al., 2022). Therefore, a key challenge for interpersonal communication is how to select the most appropriate utterance from the mental lexicon to convey communicative intent with clarity and precision.

Traditional psycholinguistic theories, such as the cohort-based, selection-by-competition model, have explained the issue of how individuals select lexical items during language production or recognition (Levelt, Roelofs, & Meyer, 1999; Marslen-Wilson, 1987). According to these theories, a set of candidate lexical forms is activated and the most contextually appropriate one is selected through competition among candidates in activation strength. While originally developed for individual-level processing, these theories may also offer insights into the utterance selection mechanism in two-way communication, where mutual understanding depends on real-time adaptation.

This study seeks to clarify the cognitive and neural processes that support utterance selection, situated within the cohort-based, selection-by-competition framework, by comparing two candidate mechanisms: statistical learning (SL) and pragmatic reasoning (PR). SL theory posits that utterance selection is determined by the bottom-up sensitivity to statistical distributional regularities of words, as evidenced by findings that infants can extract statistical patterns from speech input to support phonetic categorization and word segmentation (Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996). In contrast, PR theory emphasizes the cooperative nature of communication (Grice, 1975), proposing that speakers engage in top-down inference about their partners mental state, to optimize uttance selection (Goodman & Frank, 2016). In this study, we aimed to determine whether utterance selection is governed primarily by statistical learning or pragmatic reasoning mechanisms.

Methods

Participants

Totally, 60 healthy adults (30 females, age = 20.533 ± 2.318) were recruited to perform a collaborative word generation task. All participants were right-handed and had normal or corrected-to-normal vision. None of them have any neurological or psychiatric disorders. Written informed consent was received from all participants.

Task procedure

The collaborative word generation task, adapted from Salazar et al. (2021), required dyads to simultaneously produce identical words across 11 consecutive trials (Figure 2a). In the first trial, participants had 5 seconds to think and 3 seconds to speak, with their speech recorded via a table-mounted microphone. From the second trial onward, each audio of the voice recorded from one member of the dyad during the last trial was played back to the other member of the same dyad using an in-earphone immediately before she/he was thinking in the current trial, based on which she/he could predict the word that her/his partner might produce and then decide what to say by her/himself in the current trial during thinking. Participants self-paced the thinking period, ending it by pressing "p" or "q", while listening and speaking phases were fixed at 3 seconds each. In earlier trials, words produced by the dyad might be different. But in later trials, we expected that the semantics of words would become more and more similar and eventually reach a precise consistency (i.e., producing the same words). This procedure was repeated across four sessions, separated by 15-second resting-state intervals as baselines. fNIRS was employed to record brain activity from both participants throughout the task.

Computational modeling

The computational modeling comprised two parts. First, to probe pragmatic reasoning in collaborative word generation task, seven interpersonal prediction models were built to test how interpersonal prediction (IP) and prediction error (IPE) shape verbal responses (Figure 2d). Second, six cohort-based word selection models were built (Figure 1): one random baseline, two SL models, and three PR models. SL1 relied on word frequency, while SL2 utilized GPT-2-derived transition probabilities between context and candidate words. PR models were based on the out-of-sample reasoning mechanism derived from the optimal IP model (M4 in Figure 2e), validated at cognitive, neural, and combined levels.

Results

Behavioral performance

The number of sessions that satisfied the task requirement was counted. The results indicated, across all dyads, 56.9% of the sessions (N = 120) met the task criterion (Figure 2b). Additionally, A slope test indicated a significant increase in semantic similarity over trials ($\beta = 0.216$, $t(818) = 13.034$, $p <$

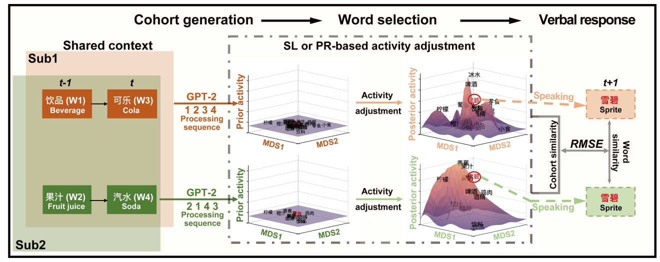

Figure 1: The structure of cohort-based word selection model.

0.001, Cohen's $d = 2.42$; Figure 2c), demonstrating robust semantic convergence.

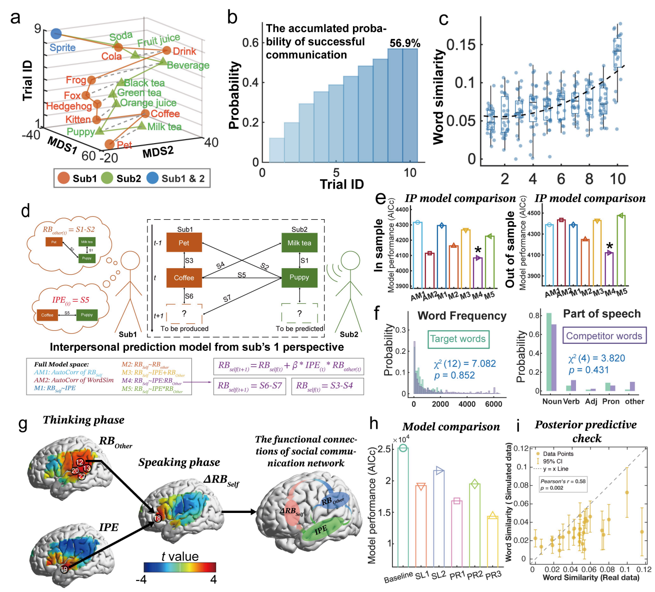

Figure 2: a) An example of the word chains from a dyad. b) The accumulated probability of successful communication across trials. c) The dynamic change of semantic similarity across trials. d) The structure of IP model. e) The IP model comparison. f) The frequency and POS distribution between human- and GPT-2-generated words. g) The neural basis of IP model. h) The cohort model comparison results. i) The posterior predictive check of PR3 model.

Interpersonal prediction model

Model comparison based on corrected Akaike Information Criterion (AICc) showed that Model 4 outperformed all others at the group level (Figure 2e), suggesting that IP and IPE jointly influence verbal response. At brain level, we observed that the left anterior temporal cortex (aTC) and the temporoparietal junction (TPJ) encode IPE and IP in thinking period, respectively ($p < 0.05$, FDR corrected at channel level, Figure 2g). During verbal production, the left inferior frontal cortex (IFC) integrates information from the aTC and the TPJ during the thinking phase, with a significant effect ($\beta = 0.114$,

$t(54) = 2.849,\ p = 0.022$, Cohen's $d = 0.384$).

LLM-based cohort generation

The GPT-2 was employed to simulate human word cohort generation. The consistence between human- and GPT-2-generated words were compare in terms of frequency and part of speech (POS). The results showed no significant differences in these linguistic features (Word frequency: $\chi^2(12) = 7.082,\ p = 0.852$; POS: $\chi^2(24) = 3.823,\ p = 0.430$; Figure 2f).

Cohort-based word selection model performance

Among six cohort-based word selection models (1 baseline, 2 SL, 3 PR, Figure 1), the PR3 model incorporating both cognitive and neural PR components achieved the best fit (Figure 2h). Posterior predictive checks revealed a significant correlation between observed and predicted semantic similarity at the dyad level ($r(24) = 0.579,\ p = 0.002$, Figure 2i).

Discussion

Resolving uncertainty in utterance selection remains a key challenge in cognitive science and LLM research. Our findings show that humans effectively manage such uncertainty in open-ended, two-way communication. Notably, statistical learning generates candidate lexical cohorts, while pragmatic reasoning guides final word selection. As natural communication typically unfolds at larger linguistic levels, future work should extend this framework beyond single-word choices to sentences or conversational turns.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. 210110393) and the Interdisciplinary Research Fundation for Doctoral Candidates of Beijing Normal University (Grant BNUXKJC2412)

References

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Grice, P. (1975). Logic of conversation. In M. Ezcurdia & R. J. Stainton (Eds.), *The semantics-pragmatics boundary in philosophy.* Peterborough, ON, Canada: Broadview Press.

Hawkins, R. D., Franke, M., Frank, M. C., Goldberg, A. E., Smith, K., Griffiths, T. L., & Goodman, N. D. (2022). From partners to populations: A hierarchical bayesian account of coordination and convention. *Psychological Review*, *130*(4). doi: 10.1037/rev0000348

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102. doi: 10.1016/0010-0277(87)90005-9

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), 101–111. doi: 10.1016/S0010-0277(01)00157-3

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Salazar, M., Shaw, D. J., Gajdo, M., Mare, R., Czekóová, K., Mikl, M., & Brázdil, M. (2021). You took the words right out of my mouth: Dual-fmri reveals intra- and inter-personal neural processes supporting verbal interaction. *NeuroImage*, *228*. (December 2020) doi: 10.1016/j.neuroimage.2020.117697