Interactive Communicative Inference

Larry Muhlstein (lmuhlstein@ucsd.edu)
University of California San Diego
Department of Cognitive Science, 9500 Gilman Drive
San Diego, CA 92093 USA

Abstract

In the search for an understanding of human communication, researchers often try to isolate listener and speaker roles and study them separately. Others claim that it is the intertwinedness of these roles that makes human communication special. This close relationship between listener and speaker has been characterized by concepts such as common ground, backchanneling, and alignment, but they are only part of the picture. Underlying these processes, there must be a mechanism for making inferences about our interlocutors' understanding of words and gestures that allows us to communicate robustly and efficiently without assuming that we take the same words to have the same meaning. In this paper, I explore this relationship between language and concepts and propose an interactive mechanism that can facilitate these latent conceptual inferences. Finally, I show how this proposal paves the way for a more precise account of the role of interaction in communication.

Keywords: Communication; Coordination; Interaction; Pragmatics; Bayes; Cognitive Linguistics; Inference; Discourse

Introduction

FRIEDA: "I study cognitive science"

FRED: "Cool! The brain is so interesting!"

FRIEDA: "Uh, cognitive science $\neq$ neuroscience..."

Human communication is fraught with misunderstanding, incorrect assumptions, and uncertainty, yet we still manage to make it work. To handle these impediments, we make ample use of processes such as grounding (H. H. Clark & Brennan, 1991), alignment (Pickering & Garrod, 2004), repair, and backchanneling (Schegloff, Jefferson, & Sacks, 1977), all of which are well described in the scientific literature. Because these processes are thought to be somewhat modular, many linguists–especially computationalists–find it useful to remove these processes from their models and experiments and reasonably assume that they can be added back in later when a more complex and complete theory is desired. While it is reasonable to excise details tangential to the core phenomenon of study, to remove these interactive processes underestimates the degree to which they are embedded in human communication and fails to appreciate how indispensable they are to communicative success.

The misunderstanding at hand emerges from the fact that these processes are not fully understood. On one hand, grounding, alignment, repair, and backchanneling allow us to repair inferential errors and to establish tighter conventions, but each of these processes also presupposes the ability to detect misalignments and miscommunications in the first place. While this might not immediately appear to be much of a problem, there is a complete dearth of empirical literature on such inferences and of any theoretical analysis about their mechanics. The purpose of this paper is to provide a preliminary analysis of the necessary and sufficient properties of these interactive communicative processes and to argue that any reasonably sophisticated understanding of human communication must build upon an epistemologically-sensitive theory of how we detect and repair misalignments and misinferences in order to communicate robustly and veridically.

Dead reckoning behavior in models of communication

Contemporary models of human communication tend to operate with strong assumptions about what I will call *conceptual alignment*–the extent to which interlocutors' mappings from surface structure (words, phrases, and actions) to hidden structure (meanings and concepts) align with one another. Conceptual alignment captures how the correspondence[1] between a speaker's beliefs about the listener's understanding of a situation and the listener's actual understanding affects their ability to communicate[2]. Typically, these models assume that both interlocutors have complete probabilistic conceptual alignment, and therefore that the listener's belief distribution about the meanings the speaker intends to communicate with her utterances is equivalent to the speaker's probability of producing those utterances given her communicative intentions. With this assumption in place, a rich subset of pragmatic behavior including all sorts of implicatures can be fruitfully investigated, and have been with the advent of Bayesian models of pragmatics such as RSA and its many variants (Degen, Tessler, & Goodman, 2015; Frank & Goodman, 2012; Goodman & Lassiter, 2014; Kao, Wu, Bergen, & Goodman, 2014).

However, if we were to drop this assumption from these models in favor of a more realistic amount of uncertainty, we would notice some problems. For example, if we allow there to be misalignment between the interlocutors' concepts, we would find that any attempt at communication results in a systematic pattern of errors that are not correctable within the scope of the model. A plausible case of this might be if I said the word "justice" intending it to mean something akin to Rawlsian fairness, where all decisions are made from a position of agnosticism about where one falls in society, but my interlocutor thought I was talking about retributive justice, then everything that I said would be misinterpreted, and this would continue indefinitely in the aformentioned mod-

[1] In a roughly Bayesian sense

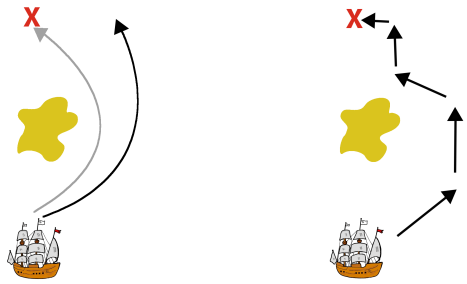
[2] Though the construction of a privileged and precise version of this concept is of considerable value, it is regrettably beyond the scope of the present article.

els because they don't contain any mechanism for detecting conceptual misalignment (instead they assume alignment and proceed from there). An even simpler case could involve a sense/reference mismatch where I use the word "speaker" to refer to a lecturer, but my interlocutor instead interprets it to refer to a sound production device. While it is likely that the vast difference in meaning would cause me or my interlocutor to notice the misinterpretation, a large class of current formal models of conversation would proceed as if I was being properly interpreted at all times.

It is useful to refer to this type of behavior as *dead reckoning communication*, as it offers the same perils as nautical navigation without taking periodic measurements of location. A ship without instruments, sailing perpetually into the fog, is inevitably bound to stray far from its destination no matter how precisely it was pointed at the beginning of its journey. This happens because of the pervasive uncertainty about the ship's motion, the effects of navigational actions (trimming the sails, hoisting the spinnaker, turning the rudder, etc.), and the environment–which is always changing. Even if a ship's captain plans out a series of actions in advance in order to get the ship to its intended destination and executes these actions flawlessly (without taking interim measurements about the position, speed, and heading of the ship), there is but a tiny chance that the ship will end up in its intended port. So it goes with communication sans feedback. If a speaker wants to convey some concept or scenario to a reader and she develops a series of communicative actions (a communication plan) and executes it without observing the listener's interpretations, then she risks the listener's gross misinterpretation unless she is willing to put in substantially more detail and effort than is typically prudent or even possible in a conversation. Even for the most closely aligned concepts, communication is bound to stray off course if the participants don't continuously probe the state of the discourse and correct its course when necessary. As H. H. Clark and Wilkes-Gibbs (1986) argued, the constraints of conversation restrict us to brief, ad hoc, ephemeral communicative actions, which limits interlocutors' ability to provide the kinds of lengthy descriptions that might be present (and necessary) in a book. However, conversation also affords interaction. As we will see, this offers additional opportunities for coordinated communication between interlocutors by allowing them to make inferences about each others' interpretations. Just like a navigator's instruments allow her to detect the ship's position and velocity and make informed corrections to its course, a speaker's inferences allow her to probe the state of her partner's understanding and choose her successive communicative actions accordingly.

Response-based inference

To combat this undesirable dead reckoning behavior, we can look both at the necessary properties for any mechanism to handle pervasive uncertainty and misalignment (given the right abstraction of the problem) and at empirical investigations about how humans specifically seem to handle it in the case of communication.



(a) Dead reckoning navigation (b) Navigation with instruments

Figure 1: Dead reckoning vs. instrument-based navigation on a ship. (a) Dead reckoning cannot handle uncertainty and so the ship ends up far from its intended destination. The gray line represents the planned trajectory, while the black line shows the actual trajectory under dead reckoning. (b) Instrument-based navigation allows the captain to correct the ship's course as it goes leading the ship exactly to its intended destination.

Challenge-response authentication

A particularly good formal analogy for the solution I will soon propose is the cryptographic concept of challenge-response authentication. Challenge-response authentication was developed as a solution to a problem much like the issue of dead reckoning in communication. Imagine that there are two interlocutors, Alice and Bob, each sitting at one end of a digital communication channel (imagine them sitting at desks in separate buildings connected via the internet.) Bob wants to send Alice a secure message, but first Bob needs to know for sure that he is communicating with Alice and not with a malevolent interloper such as their friend Charlie. To verify Alice's identity, Bob needs Alice to say something that only she would know, such as the secret phrase they agreed upon earlier. Bob receives this phrase from the communication channel and concludes that it must have come from Alice. But Charlie, the conniving and generally clever chap that he is, has tapped the communication channel and observes Alice and Bob's secret phrase. The next day, when Bob wants to communicate with Alice, he hears this same phrase, and again believes that it came from Alice. Unbeknownst to him, Alice was sick and wasn't at the computer terminal she normally uses to communicate with Bob. Instead, the secret phrase came from Charlie, who now proceeds to communicate with Bob as if he was Alice.

A natural question then follows: how can Bob and Alice communicate securely without risk of Charlie impersonating Alice? The answer lies in challenge-response authentication. Bob, having discovered the fatal flaw in his authentication system, comes up with a clever alternative. Instead of agree-

ing with Alice upon a secret phrase, they instead come up with a secret relationship between phrases–i.e. a secret function. The next day, when Alice and Bob try to communicate, Bob sends Alice a message called a *challenge*. When Alice receives this challenge, she feeds it into the secret function and gets an output that she then sends back to Bob as a *response*. When Bob receives this response, he feeds his challenge into his own copy of the secret function and compares the result with Alice's response. If they match, then Bob can be sure that he is communicating with Alice and not Charlie. To understand why this works, we can look at Charlie's behavior in this scenario. Like before, Charlie has tapped the communication line and receives both Bob's challenge and Alice's response. The next day, when Charlie tries to pretend that he is Alice, he receives a challenge from Bob. This challenge however, is not the same as the one he saw the day before. This means that the correct response is different as well! Because the secret knowledge is a full function, the observation of a few challenge-response pairs is not sufficient to induce the full function or the correct response to additional challenges. If the secret functions are properly designed, then virtually no amount of observations of the challenge-response pairs will be sufficient for Charlie to induce the secret function. Upon realizing this information, the sullen Charlie decides to leave Alice and Bob alone to search for others with weaker verification algorithms to deceive.

Utterance-response contingencies

How does this cryptographic mechanism relate to human communication and how does it help us avoid dead reckoning behavior? The problem in human communication is not in establishing the identity of the interlocutor, but rather in verifying the interlocutor's comprehension. If we replace the secret functions from challenge-response authentication with out latent conceptual understandings, we can use the same sort of strategy to verify the similarity or alignment between our communicative intentions and the interlocutor's interpretations. If Alice produces utterance x for Bob, and Bob responds with utterance y, then Alice can check Bob's comprehension of x by modeling the plausibility that Bob would have generated response y given various interpretations of x. In the simple case, if Bob interprets x exactly as Alice intends, then his response y will be identical to Alice's prediction about his response given her intended interpretation. In the case where there are multiple plausible interpretations, Bob's response y provides Alice with information about how likely Bob is to have made each of the possible interpretations of x. We will call this minimal pair of utterance and response an *utterance-response contingency*. Utterance-response contingencies are the basic building blocks of *interactive communicative inference*, a term I propose to denote the general process of inferring interlocutor beliefs through interaction. This also involves more complex cases where each interlocutor's beliefs about the others' understanding are updated through extended dialogical interaction, which can result in robust alignment processes that make effective discourse possible.

In real discourse, utterance-response contingencies look like this:

GEORGE (UTTERANCE): "Wasn't that a great speech yesterday?"
GEORGIA (RESPONSE): "I mean, it wasn't as horrible as I expected, but I definitely can't say it was good..."
GEORGE (REPAIR): "Oh, did you think I meant Donald's speech? I was talking about the address from the director of the ACLU."

These utterance-response contingencies can come in various forms. They can resolve referential misinterpretations like in the example above, or they can surface more subtle conceptual misalignments where the interlocutors' word-concept mappings are misaligned, or even when the internal structures of their respective concepts are inconsistent with each other[3].

The case of misaligned word-concept mappings is illustrated in the opening dialogue between Frieda and Fred.

FRIEDA: "I study cognitive science"
FRED: "Cool! The brain is so interesting!"
FRIEDA: "Uh, cognitive science $\neq$ neuroscience..."

Here we see that Frieda's concept of "cognitive science" is reflective of her being an insider to the field and therefore likely includes associations with each of the "six corners of the hexagon": psychology, computer science, philosophy, linguistics, anthropology, and neuroscience. Fred's concept, however, is closer to how many outsiders think of cognitive science[4]–as alternative word for neuroscience. Using interactive communicative inference and utterance-response contingencies, Frieda is able to notice this conceptual misalignment and repair it, thus improving the alignment between interlocutors.

A more systematic formulation

To make this idea a bit more concrete, we can situate it in the context of a more precise way of looking at human communication. Fundamentally, communication is comprised of two kinds of processes: inferences, and actions that facilitate inferences. The idea of an *inference* is mostly self explanatory. It refers to any kind of inference about the beliefs, communicative intentions, or mental state of your interlocutor. Such beliefs can be about things in the world or more abstract concepts, and so the contents of these inferences are almost unlimited in scope. *Inference facilitation*, on the other hand, refers to actions that are taken by one participant in the discourse (whom we call the speaker, even though their actions may not be vocally produced) and observed by another participant in the discourse. As these actions are caused (at least in part) by the speaker, they provide the observer of the actions (the listener) with information

[3]Such as in the earlier example involving misaligned concepts of justice.

[4]At least in this author's experience

about the speaker and consequently their beliefs, mental state, and communicative intentions. Since the roles of speaker and listener shift dynamically throughout a discourse, we do not associate these terms with particular conversational agents, but with the roles themselves. The listener–in responding to the speaker–therefore temporarily instantiates the role of the speaker, and even the most subtle facial expressions, when produced in a discourse, count as inference-facilitating actions. Besides these two components, there is also the idea of *context*, which is broadly defined to capture any effect of the discourse topic, the surroundings, or any other processes on the particular instance of communication[5]. The definitions of these communicative concepts are not constrained further than this because to do so would limit their generality as the basic constituents of human communication.

I choose this framework instead of other possibilities such as the pragmatic alternatives framework, classical communication theory, and informal schools of thought such as relevance theory and cognitive linguistics because it affords both precision and generality and it allows us to highlight the basic epistemics of human communication. In this formulation, all communicative actions have the same status. This is because, at the epistemic level, an utterance, a gesture, and an unintentional twist of the lips are all actions taken by a communicator that allow an interlocutor to make various inferences about her communicative intentions and her state of mind. While the contents and causal pathways associated with these actions may differ drastically, these details do not have bearing on the development of a basic framework for communicative analysis, only on specific theories subsequently derived from that framework.

To move towards a formal theory, we can choose a minimal Bayesian framework that captures only the epistemic relationships between the speaker's utterance, the listener's interpretation, the listener's response, and the context. Since all of these, except for the listener's interpretation, are observable to both parties[6], we can capture their relationships in a causally-derived probabilistic graphical model where each of the nodes is observed except for the node representing the listener's interpretation, as seen in figure 2 (Pearl, 1988). This model captures the fundamental structure of the speaker's inference about the listener's interpretation via utterance-response contingencies, where the details of particular inferences depend on the particular distributions that comprise an instantiation of the model. While the broad epistemic structure of these inferences is simple, the corresponding real-world processes are anything but so. Filling in the details of this model will not be an easy process and will require a large amount of directed scientific experimentation and theoretical analysis. Therefore the goal here is not to develop a full theory of these communicative processes in real humans, but to provide a structured framework through which we can understand and investigate them. Bayesian probability is sufficiently abstract to allow us to represent these general inferential structures while allowing the rich human details to be added later.

[5]Context is an incredibly complex topic, but for the purposes of this article, we will leave its nuances aside.

[6]Except possibly some component of the context, but the variance in context between participants in an interaction is beyond the scope of this article.

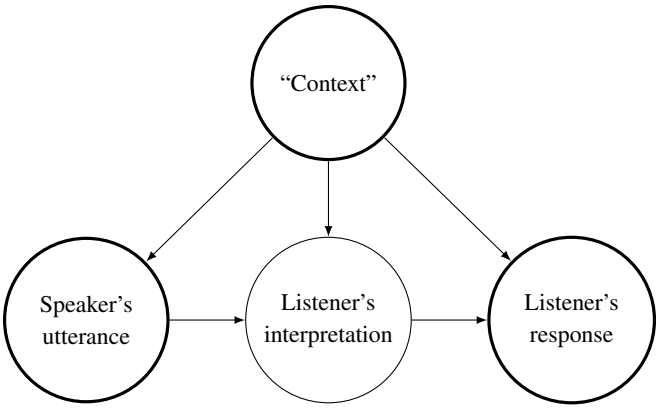


Figure 2: Bayesian network representation of interactive communicative inference. Bolded nodes are directly observable.

Application to known discourse phenomena

To further ground this framework and illustrate its relationship to empirical phenomena in human communication, we will look at three different concepts described in the literature and show how the idea of interactive communicative inference offers each of them a stronger theoretical foundation.

Common ground

The idea of common ground, first proposed by H. H. Clark and Brennan (1991), suggests that people accumulate a shared repository of knowledge when they interact, and that subsequent interactions are facilitated by this common knowledge. This proposal has received a substantial amount of theoretical analysis, which has lead to a rich account of how humans establish common ground and make use of it in conversation (H. Clark, 1996). There is also a solid foundation of experimental evidence supporting the theory (Brennan & Clark, 1996; Hanna, Tanenhaus, & Trueswell, 2003). Perhaps the most notable experiment–known as the "tangrams experiment"–involved a pair of interlocutors who were given a set of cards depicting images of blocky figures–i.e. tangrams–and were tasked with getting the other person to arrange their cards in an order perceivable only to the designated speaker for the round. Complicating this significantly was the fact that the participants were separated by a visual wall and so could only communicate verbally. H. H. Clark and Wilkes-Gibbs (1986) found that the participants quickly established shared conventions for referring to the cards, which resulted in a decrease in the amount of communication required to complete the task as they continued to interact through multiple task rounds.

While this story is compelling both theoretically and empirically, it is missing a mechanistic account of the inference

processes that these interlocutors go through when building and making use of common ground. Dead reckoning models cannot capture this progressive coordination behavior, and so we turn to interactive communicative inference. The basic component of interactive communicative inference, the utterance-response contingency, accounts for how an exchange can allow the speaker to update their understanding of the listener's beliefs and therefore facilitate the listener's comprehension. When this occurs, the speaker can choose to use a referring term that they know the listener will understand, which allows the speaker to communicate more efficiently with the listener via this ad hoc convention. If both communication partners make use of such processes in their interaction, then they will come to a shared understanding of how to use language to communicate particular meanings. Over an extended interaction, such interlocutors can build a common communication system. This predicts the increasing communicative concision throughout an interaction that H. H. Clark and Wilkes-Gibbs (1986) observed. We can even imagine that, if this process involves additional people over multiple interactions, a complete conventionalized communication system should emerge from this process of building shared knowledge via interaction.

Conversational alignment

Martin Pickering and Simon Garrod have developed a compelling theory of discourse as an interactive alignment process, which has a number of useful relations to the present theory of interactive communicative inference (Pickering & Garrod, 2006, 2004; Garrod & Pickering, 2004). Their theory claims that local lexico-syntactic priming in discourse produces a series of cascading effects that causes all participants in a conversation to produce similar surface structures and even to align on semantic content. Using this theory, they argue that the fundamental mechanisms operating in discourse are these alignment processes and construct an account of language processing that does not require a speaker to maintain an explicit and complex model of their interlocutor (Garrod & Pickering, 2004). While this theory accounts for a wide range of phenomena, it does not offer any account of conceptual alignment phenomena. Conceptual alignment cannot arise from priming-based mechanisms because it is defined as the alignment between latent conceptual structures and surface communicative actions, and this relationship cannot be primed by observing surface structures from an interlocutor's communicative actions. In order to account for the facts that our conceptual structures tend to be relatively aligned within a conventionalized communication system and that we have evidence that they are aligned via interaction, an interactive communicative inference component needs to be added to Pickering and Garrod's theory of discourse.

Their theory also predicts that it should be easier for humans to efficiently communicate meaning through interactive dialog than unidirectional monologue because interaction facilitates the alignment of linguistic representations (Garrod & Pickering, 2004). This prediction is both consistent with our informal experience as well as predicted by the interactive communicative inference framework. For a monological description to yield effective communication, it must necessarily contain enough information such that, for all of the likely ways that a listener or reader might misinterpret the meaning, there is additional information that steers them away from these misinterpretations and towards the intended interpretation. In a discourse, however, these counterfactuals do not need to be handled via anticipation and mitigation. Instead, utterance-response contingencies make it so that only actual misinterpretations by the listener need to be ameliorated.

With the addition of the present account of conceptual alignment through interactive communicative inference, we can add a vital component to the conversational alignment story to yield a powerful framework for understanding human communication.

Backchanneling and repair

Sociolinguistics and conversation analysis research has built an account of discourse expressed in terms of the constituent behaviors of conversations and the types of communications they enable (Sacks, Schegloff, & Jefferson, 1974). Some of the core concepts in these accounts are the ideas of backchanneling and repair. *Backchanneling* is invoked in a multichannel model of communication in which interlocutors communicate the majority of their content via a main channel and provide meta-conversational signals in a backchannel (Yngve, 1970). For example, a listener may produce the affirmative "uh huh" in a backchannel to signal to the speaker that the she believes herself to be comprehending and that the speaker should proceed. While these behaviors have been convincingly shown to play a key role in natural conversation, this account leaves open the question of how the listener forms her beliefs about whether or not she is comprehending the speaker. As we have seen, belief of comprehension does not necessarily imply veridical comprehension, because there may be an undetected conceptual misalignment between the speaker and listener. Our account of interactive communicative inference suggests that listeners may be providing these cues to speakers in order to facilitate the speakers' interactive communicative inferences by completing the utterance-response contingency. It also suggests that speakers may be comprehending these backchannel signals, not truly as an unequivocal signal to proceed, but as a response that conveys the listener's beliefs about their comprehension. The following dialog illustrates a scenario with a distinction between the affirmative backchannel as an unequivocal signal of comprehension and an account of backchanneling as concurrent response. (Tyra and Tyler are roboticists at different universities.)

TYRA (UTTERANCE): "STEVE can walk on two legs now!"

TYLER (BACKCHANNEL): "Oh"

TYRA (REPAIR): "Not Steven the child, STEVE (Self-Taught-EVacuative-Entity) my robot. It's a ground-

breaking achievement!"

Here Tyra infers Tyler's miscomprehension through his affectless backchannel response to her achievement. She knows that he would be really excited to hear about this research achievement and therefore infers that he must believe her to be talking about a child instead. Tyler mistakenly believes himself to have comprehended Tyra's utterance and produces a backchannel response, but Tyra does not interpret it blindly. Instead, she uses the resultant utterance-response contingency to infer the miscommunication and repair the misunderstanding. As predicted in the classical account, the backchannel provides feedback to Tyra about Tyler's interpretation. However, the response does not reflect Tyler's veridical comprehension, but facilitates a more complex interactive communicative inference.

Miscommunications in conversations are corrected via *repairs*. These come in many forms, but are often divided into two classes: self-initiated and other-initiated. While the literature provides substantial experimental and ethnographic detail about the role of repairs in conversation, it does not offer an account of how interlocutors infer when a repair needs to be made (Schegloff et al., 1977). The present theory of interactive communicative inference offers a computational mechanism by which miscommunications can be detected in conversation, which allows speakers to repair the miscommunication by correcting their own production or by correcting the listener's interpretation.

Conclusion

Contemporary models of communication are incomplete. While they offer a detailed understanding of the surface-level phenomena present in discourse, they do not provide a satisfactory explanation of the inferential mechanisms necessary for these phenomena. They are unable to account for the robustness of communication in spite of uncertainty, how people know when to update their beliefs about their interlocutors, and how people establish conventions. I have provided theoretical evidence tied to the empirical to show how all of these types of missing accounts can be derived from the idea of interactive communicative inference.

While researchers have long understood that discourse is not a unidirectional and isolated activity, we have demonstrated that treating it as such, even for the sake of delimiting the domain of a theory or a model, can have the effect of removing an important property of language. When the idea of interactive communicative inference is taken seriously, we can begin to construct a scientific study of communication that can account for how, despite the fact that we can never be in someone else's head to see first-hand what they believe and how they feel, they are able to show us these things simply by engaging us in a cooperative dance of action and interpretation.

Acknowledgments

I would like to thank Sean Trott for helping motivate me to write this article as well as for his invaluable feedback. I would also like to thank Dorothy Howard for her time helping me polish the article. Finally, I am deeply indebted to the UCSD Department of Cognitive Science and all of its members for their continued support.

References

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482.

Clark, H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on socially shared cognition*, 127–149.

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

Degen, J., Tessler, M. H., & Goodman, N. D. (2015). Wonky worlds: Listeners revise world knowledge when utterances are odd..

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in cognitive sciences*, *8*(1), 8–11.

Goodman, N. D., & Lassiter, D. (2014). Probabilistic semantics and pragmatics: Uncertainty in language and thought. *Handbook of Contemporary Semantic Theory. Wiley-Blackwell*.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, *49*(1), 43–61.

Kao, J. T., Wu, J. Y., Bergen, L., & Goodman, N. D. (2014). Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences*, *111*(33), 12002–12007.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(02), 169–190.

Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*(2-3), 203–228.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *language*, 696–735.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 361–382.

Yngve, V. H. (1970). On getting a word in edgewise. In *Chicago linguistics society, 6th meeting* (pp. 567–578).