Joint epistemic engineering: The neglected process of context construction in human communication

Arjen Stolk,[1,2] Jana Bašnáková,[1] Ivan Toni[1]

[1]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, the Netherlands

[2]Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

Abstract

This contribution argues that a common language and its statistics do not explain how people overcome fundamental communicative obstacles. We introduce joint epistemic engineering, a neurosemiotic account of how asymmetric interlocutors can communicate effectively despite using signals that are referentially contingent on the current communicative circumstances. The basic insight is that a communicative signal contains a multiplicity of functions and that interlocutors use those multi-layered signals to simultaneously coordinate a space of possible interpretations, declare a communicative intent, and to reduce uncertainty over the identity of a referent.

Keywords

Social interaction, mutual understanding, conceptual alignment, brain, autism

"There is not much dependence to be placed upon these Constructions that we put upon Signs and Words, which we understand but very little of, & at best can only give a probable Guess at their Meaning."

-- David Samwell, ship surgeon on James Cook's HMS Discovery, Hawaii, 1779

There is virtually no limit to what human minds can conceive beyond the physical world. Those mental activities might appear inconsequential or delusional. Yet when people collectively agree on mental constructs, those cultural innovations can change reality, as when we attribute value to intrinsically worthless banknotes. Ultimately, those cultural innovations are rooted in our communicative abilities, but it remains a mystery how individual minds, each with personal backgrounds and perspectives, converge on seemingly similar mental constructs in communication (Corballis, 2017; Levinson, 2006; Misyak et al., 2014). This contribution focuses on the neurosemiotic mechanisms supporting this conceptual alignment challenge, and illustrates how a largely neglected epistemic engineering process permeates human communication.

A core assumption in the language and communication sciences is that conventional behaviors such as words, gestural emblems, and facial expressions constitute the primary vehicle for transferring information between interlocutors. Formal accounts have suggested that some of those behaviors have been optimized to convey meaning reliably, with a pre-defined syntax combining their meanings compositionally to form a larger meaning (Gibson et al., 2019). Indeed, examination of information transfer through a lens of systematicity has revealed several regularities in word order and turn-taking patterns across languages and conversations (Dryer, 1991; Kendrick & Drew, 2016; Sacks et al., 1974). These stereotyped dependencies between words and dialogic turns arguably help interlocutors by providing a predictable shared structure and cues to a meaning. However, stereotyped dependencies alone do not give the full meaning of an utterance, otherwise people would take the sentence "John dressed and had a bath" to unequivocally mean that John took a bath with his clothes on, which they rarely do (Fillenbaum, 1974). This is not a glitch in an otherwise flawless system (Clark, 1997). Everyday interaction is filled with ambiguity and implicature, not to mention vagueness and the occasional misdirection or deception. Everyday interaction is also structurally "messy". An utterance can constitute a response to another verbal or non-verbal utterance that occurred at an arbitrary point in time along the interaction's trajectory, thus irrespective of word order, turn sequence, or communication modality. These flexible conceptual dependencies between utterances constitute the bulk of our daily communications and cannot be easily analyzed outside the communicative contexts in which they are embedded, as the virtual assistants on our phones try to do (e.g., Alexa, Siri, Google Assistant). This contribution articulates a number of outstanding issues in the study of human communication, and reviews recent work in neurosemiotics that has begun to shift the focus from context-free information transfer to the construal of meaning within genuinely interactive contexts. We provide arguments on why the field can and should move beyond views of human communication as information transfer, where signals have stereotyped and publicly invariant consequences. The real challenge is to specify how people in dialogue define a context rich enough to afford a signal to have a fair chance at being interpreted in the way it is intended at the moment of delivery, yet flexible enough to be continuously adjusted to the goals and needs of the conversation. We elaborate on recent contributions highlighting the importance of a shared conceptual space in human communication (Stolk et al., 2016; Toni & Stolk, 2019), delineating possible mechanisms leading to the joint construction of that knowledge space. We argue that a large portion of the context of use of a signal is constructed on-the-fly by members of the interaction, a process we call "joint epistemic engineering". In what follows, we illustrate how interlocutors use multi-layered communicative signals as a tool to mutually align and update their conceptual structures during social interaction, providing a critical and situated source of interpretational constraints. By embedding signals in the shared conceptual space generated through this engineering process, interlocutors rapidly resolve the multiple ambiguities in every signal, coordinate current reference points, and prepare the ground for anchoring new reference points in their conceptual frame of reference (Clark, 2020).

From pre-Darwinian types to contingently shared tokens

By way of analogy, the focus of current research on communication as information transfer is reminiscent of the focus that pre-Darwinian natural sciences had on biological types. In that pre-Darwinian view, time-invariant species were the fundamental units of biological organization, worthy of investigation because of their presumed immutability and optimality (Mayr, 1999). Arguably, one of Darwin's main conceptual innovations was to shift the focus of inquiry from biological types to tokens, i.e. from ideal Platonic species to individual organisms that might or might not pass through the filter of natural selection. In Darwin's framework, each accident of history became informative, showing how a replicator adapts to the current environmental circumstances in light of past adaptations. Notwithstanding this paradigm shift in the life sciences, the focus of language and communication research has largely remained anchored on signal types, as seen for instance in the academic space devoted to investigating regularities in word order patterns (Gaskell, 2007). Accidental variation in those regularities, frequently occurring in real-life dialogue (Stivers, 2010), have been largely ignored as stochastic phenomena, rather than informative objects of study able to reveal how interlocutors create and adapt to the current communicative circumstances.

The premise of this pre-Darwinian focus is that linguistic types offer optimal solutions to the problem of linking a set of signals to their referents, with a pre-defined syntax combining referents compositionally to form a larger meaning (Gibson et al., 2019). Those solutions are optimal insofar human communication can be framed as an information-theoretic problem. In that framework, the main objective is to reduce uncertainty over the identity of a referent brought about by the detection of a signal (Shannon, 1948). However, without pre-existing and (largely) overlapping signal-referent mappings, a signal cannot reduce the uncertainty of a receiver on the sender's intended referent of that signal. In daily language use, those mappings are rarely there beforehand, nor are they overlapping across interlocutors. We tend to take for granted that linguistic and extralinguistic signals like words and emblems have meanings shared with other members of our linguistic community, up to the point where those meanings are considered properties of those signals. This fallacy of symbolic interpretation is analogous to the misattribution of a color-percept exclusively to the spectral distribution of light bouncing off an object, when in fact a color-percept is an active cognitive inference contingent on the spatial and temporal context in which that spectral distribution occurs (Conway, 2012). As a painter needs to decide which hue to make from her palette for achieving a fair chance the viewer will perceive a red apple in that particular canvas (Lotto & Purves, 2000), so a communicator needs to decide how to make a signal that will be interpreted as intended in the current communicative context. Counting on context-invariant signal-referent mappings, overlapping across interlocutors, is not a viable communicative option.

Building on these considerations, this chapter raises two main conceptual issues, in line with a large body of work on experimental semiotics and pragmatics (Bašnáková et al., 2014; Galantucci & Garrod, 2011; Noveck & Reboul, 2008; Scott-Phillips, 2017). First, we advocate a shift from pre-Darwinian typological thinking towards studying the actual tokens used during communicative interactions, considering each signal as contingent on the current communicative circumstances. We argue that human communication should be thought of as joint control of a shared conceptual space, rather than signal coding-decoding. Second, we emphasize the consequences of the fact that rarely, if ever, a communicator can assume overlap with an addressee. Interlocutors' overlaps in signal-referent mappings and in communicative contexts are a product of communication, not a default state. We explore a way these interpersonal disparities can be resolved on the fly, namely the suggestion that a communicative signal contains a multiplicity of functions. We argue that a communicative token, besides being deployed to reduce uncertainty over the identity of its referent (a Shannon-signal, (Shannon, 1948)), is also meant to ostensively mark its own communicative value (a Grice-signal, (Grice, 1975)), and to offer cues for aligning the space of possible interpretations across interlocutors (a Peirce-signal, (Peirce, 1931)).

What is in a signal?

Shannon-signals

In the original transmission-capacity view of communication, formulated by Shannon (Shannon, 1948), there is no mechanism for assigning meaning to a signal. A signal conveys information to a receiver when it reduces the receiver's uncertainty over the identity of that signal given the set of possible signals and referents. It is assumed that both sender and receiver share a pre-existing set of signal-referent mappings, like two telegraph operators sharing the Morse code. Natural and simulated selection has led to numerous situations where conspecifics can rely on shared signal-referent mappings (Danchin et al., 2004; Kirby et al., 2014). In those situations, a communicator can generate Shannon-signals, i.e. signals with fully predictable and stereotyped referents, since the set of possible signals and the set of signal-referent mappings are already fully shared by senders and receivers. Alarm calls of vervet monkeys, bees' dances, or human behavioral contagion are just a few examples of Shannon-signals. However, this communicative scheme requires repeated rounds of natural selection under consistent pressures to arise. Accordingly, it is not clear how it can account for the referential flexibility of human communication, as seen in the moment-to-moment variability in signal use across and within conversations (Brennan & Clark, 1996). By the same token, it is also not obvious how referential flexibility can be accounted for by explanatory frameworks that ignore signal-referent mappings altogether, e.g. frameworks where signal use is determined exclusively by an iterated process of reciprocal influence and resistance between individuals, devoid of any reference (Dawkins & Krebs, 1978; Rendall & Owren, 2013).

Grice-signals

Addressing referential flexibility, Grice initiated an influential perspective on human communication, arguing that communication requires interlocutors to have knowledge about one another's intentions and beliefs in order to disambiguate the intended meaning of a signal (Grice, 1957). A Grice-signal involves a sender producing a signal to i) induce functional changes in an interpreting receiver, ii) assess that the signal has been recognized as to induce those changes, and iii) ultimately give rise to functional changes that incorporate consequences of that recognition. Put differently, a communicative signal is more than a Shannon trigger in a pre-existing signal-referent mapping scheme. As when a driver flashes her lights to an oncoming car, a signal is also meant to ostensively mark its own communicative value, signaling that we not only intend to influence the receiver (through the production of a stimulus), but also want them to recognize that we are acting with such intentions and, ultimately, respond on the basis of recognizing this. In this scheme, interlocutors can go beyond literal meaning as long as their respective speech acts adhere to the accepted purpose or direction of the conversation as well as universal principles of cooperation. As part of the latter, Grice indicates that each individual's contribution ought to be as relevant and informative as required, but not more than required for the current purpose of the exchange. However, he leaves unspecified how interlocutors determine what quantity of information is appropriate for the current purposes. What is considered concise by one, could be pedantic to another.

Peirce-signals

Emphasizing interpretation, Peirce's triadic model highlighted a special role for meaning-making agents in the communication process, adding a third entity to dyadic models of communicative tokens and the mental concepts they kindled. A Peirce-signal (he used the term of sign) is what it is interpreted to be, as when clouds signify rain to some, and ruins stand for ancient civilization to others. The triadic representation of a referent entails that a Peirce-signal can take any form, and that nothing is a communicative token unless it is interpreted as such (Peirce, 1931). Peirce further distinguished three primary signal types - icons, indices, and symbols. The former two bear a natural relationship with the object they represent, either through perceptual resemblance (icons, e.g. two fingers up for the number Two) or through physical or causal connection (indices, e.g. two fingers for V for Victory). The latter, symbols, signify an object only via an arbitrary rule imposed by a community, e.g. two fingers up conventionally standing for Peace. The problem with this tripartite scheme, as is often the case with typological thinking, is that a closer look at everyday interactions frequently yields exceptions to the rule, in this instance that utterances can be composites of two or more signal types. For example, uttering "I caught a fish this long" while holding hands a certain distance apart both symbolizes and denotes the length of a caught fish (Clark, 1997). Peirce was aware of this obvious issue, considering the types not mutually exclusive. More generally, however, until we understand how people integrate or otherwise deal with different signal types, any typological framework of human communication remains fundamentally limited. Or as the ship surgeon who sailed with James Cook noted in 1779, there is not much dependence to be placed on signal types: they only give a probable guess at a signal's meaning, raising the fundamental issue of understanding how people end up making similar guesses during communicative interactions.

Fixed-code fallacy

Overall, the multiplicity of functions and referents of a signal argue against the "fixed-code" view of signals that has percolated into modern perspectives of language and communication (Harris, 1998). In these contemporary yet fundamentally pre-Darwinian perspectives, words have universal referents within a linguistic community and encode semantic constructs of thought. This fixed-code approach is one of invariant and orderly signal-referent mappings that allow encoding and decoding of thoughts by those who know the code. Words embody cultural conventions worthy of becoming lexicalized and passed to the next generation, and their meaning is universal enough for all speakers of the language to have access to all their properties. We argue that this fixed-code view of language use is an illusion. We tend to think of language as being lawfully constituted by phonemes, words, and sentences, when in fact those elements are regularized and approximate low-dimensional projections from high-dimensional spoken utterances (Goldberg, 2019; Port, 2010). Similarly, we tend to think of communication as being constituted by signals exchanged between interlocutors, when in fact those signals are approximate low-dimensional behavioral projections from high-dimensional conceptual spaces in the interlocutors' minds. We confuse regularities in the signal (e.g. the sound /dɔg/, a low-dimensional behavioral pattern that can be produced/perceived by communicators) with regularities in the content (a referent of /dɔg/, high-dimensional concepts that require a degree of overlap across communicators to afford comparable, contextually appropriate interpretations of /dɔg/). We argue that communication is not primarily solved by interlocutors approximating the same conceptual type through exposure to a publicly available signal and a shared lexicon. The reason is that, when using language in interaction, the trajectory from a behavioral signal to its referent is anything but fixed. We argue instead that a signal-referent trajectory is contingent on the current communicative context, thus only minimally constrained by the interlocutors' individual lexica. Effective communication requires two interlocutors to invert a candidate behavior, given the current presumed context, into partially overlapping concepts. That epistemic inference, jointly engineered by the interlocutors themselves through the creation of a presumed context, allows those interlocutors to resolve on-the-fly the multiple ambiguities that permeate almost every human communicative interaction, such as distinguishing between a reflexive and an embarrassed cough, or between /dɔg/ as a pet and /dɔg/ as a villain. These considerations on the context-dependency of a signal beg the question of how interlocutors determine what counts as context, a fundamental yet largely overlooked issue we will turn to next.

What counts as context?

Disambiguation

When context is brought into models of human communication, it is to arbitrate among the possible interpretations of a signal, as when adjudicating between the meanings of "financial institution" and "riverside" for the word "bank". This view acknowledges the ambiguities intrinsic in many words. However, there are several problematic assumptions behind the notion that these ambiguities can be fully resolved through contextual factors biasing interpretation toward an intended referent. First, that disambiguation procedure presupposes largely overlapping conceptual spaces across interlocutors when interpreting that word (Warglien & Gärdenfors, 2015). Given that no two people have exactly the same experience and expertise, the fact that they are using the same expression does not guarantee the same richness of their conceptual knowledge - when a mechanic talks about a car, she thinks in terms of "brake horsepower" and "electronic stability control", a representation only marginally overlapping with how her customer conceptualizes the same vehicle. Even if expertise in the moment of the conversation was broadly comparable, it seems unlikely that the condition of perfect symmetry between interlocutors' conceptual spaces can be a priori satisfied in the longer term, given how concepts can change over a lifetime, and even within the timespan of a single conversation (Carpendale & Lewis, 2004). It is also unlikely that perfect symmetry is granted by community membership, as no community is sufficiently homogenous to provide full conceptual alignment across interlocutors (Clark & Marshall, 1981; Lewis, 1969). Second, even assuming perfect symmetry between interlocutors, context is not a clearly defined background, built bottom-up from the sensory data. Signal-referent mappings are plagued by ambiguity; contexts are at least as ambiguous, and harder to define. We do not have sensory receptors for context, e.g., a universal relevance detector (Sperber & Wilson, 1986). Third, calling upon context to achieve disambiguation ultimately moves the explanatory problem up a level, without delivering any actual explanation other than ad-hoc heuristics with no generalization power.

Strategic reasoning

It could be argued that tacitly converging on a context is a problem that has already been solved in "game theory", the study of strategic coordination among rational agents. Moving beyond the standard practice of assuming full knowledge of goals and pay-off outcomes built into a coordination game (cf. signal-referent mappings), Schelling questioned how people could tacitly coordinate given minimal cues (Schelling, 1960). For instance, study participants are given hypothetical problems like meeting with a stranger in New York City on a set date, requiring choosing a time and place. Participants tend to converge on a small set of solutions, a popular answer being noon at Grand Central Station in this particular task. Although a topic of debate, the emergence of these "focal points" have been largely attributed to strategic reasoning, either in the form of modeling and choosing the best response of other players, or of an imaginary group leader wanting to maximize group interest. These strategies would arguably afford sufficient flexibility to arrive at a different focal point (e.g., the Statue of Liberty) given another relatively well-defined situation (e.g., the stranger is a tourist). However, strategic reasoning presupposes rather than explains how rational agents come to operate on identical levels of granularity in their "New York" representation. For instance, it is unclear whether strategic reasoning would generate effective focal points across agents that do not know whether the coordination challenge is about New York City or New York state. Everyday interlocutors operating in open-ended situations cannot presume overlapping granularity in their contextual representations absent any mutual coordination.

Mutual coordination

It should follow from the above that "context" is as ambiguous and cognitively construed as the signal being disambiguated. We argue here that it is through the production and interpretation of the Peirce-signals embedded in their communicative history that interlocutors generate a shared context for constraining the possible interpretations of a signal. Besides exchanging signals, communicators jointly determine the detectability and interpretability of those signals. For instance, a communicator could furnish the sound /bæŋk/ with a pointed gaze towards the cupboard where both interlocutors know fishing equipment is stored. The mutually-known knowledge about the cupboard content, and the physical presence of that piece of furniture, offer a situated source of interpretational constraints for the /bæŋk/ sound. The production of an additional Grice-signal (the pointed gaze) poses an interpretational request at the addressee's door, aimed at aligning the interpretational frames of /bæŋk/ across the two interlocutors. The addressee might get at the intended referent if she zooms into the focal point provided by the conjunction of the cupboard and her pre-existing shared knowledge with the speaker (the fishing equipment). This trivial example of a daily interaction illustrates the inadequacy of fixed-code approaches to human communication, and the hard inferential labour required to produce and interpret a multilayered signal.

Needless to say, the Peirce-signal generated when combining the gaze, the cupboard, and the presumed mutually-known knowledge might fail to evoke in the addressee the intended context for the /bæŋk/ sound. Under those circumstances, repair operations, including asking for clarification, could prove important for achieving successful communication (Dingemanse et al., 2015). However, repair cannot be the main mechanism supporting successful communication. Calling upon repair to achieve disambiguation ultimately moves the explanatory problem into a secondary channel at least as complex as the original dialogue. In fact, repair procedures seem likely to be computationally more demanding, since the problem of figuring out what is being requested (or what to request) is compounded by the problem of providing a solution to the original ambiguities, and both are rapidly aggravated as time elapses in the turn-taking dynamics. Accordingly, it might not be too surprising that interlocutors often tolerate possible misunderstandings to avoid interrupting a dialogue (Jefferson, 1988). This tolerance towards possible misunderstanding might not even be specifically linked to disagreements. It might be a feature built into the process leading to the joint creation of a presumed context, since interlocutors need to constantly juggle multiple possible interpretational frameworks. In the example above, the addressee might try to embed /bæŋk/ into different contexts, e.g. a financial institution, a geographical feature, or a meteorological phenomenon as in a bank of clouds. The addressee might tentatively choose the financial option, but then pivot towards the geographical framework as that initially un-chosen option proves more consistent with the evidence provided by the interlocutor. Tolerance towards potential misunderstanding might result in enhanced dialogue functionality, but the driving factor of that tolerance might be the need to maintain epistemic flexibility in the face of signal-referent volatility. This suggestion implies that human communication relies on a continuous background of counterfactual reasoning (Boorman & Rushworth, 2009), contingently updated conceptual knowledge (Stolk et al., 2015), and decisions about which elements of private knowledge ought to be shared (Bang et al., 2020).

This suggestion also implies that joint epistemic engineering, differently from traditional engineering, is not geared towards predictions based on past behaviors, balancing feedback and feedforward gains to keep a system within a desired performance envelope. The irreducible uncertainty and volatility of interacting with other minds requires engineering an adaptive system that can find solutions to the problem of rapidly converging on a mental construct without knowing in advance the precise behavioral signal most conducive to those solutions. This system can tolerate rapid local changes (e.g., phonetic variation, gestural noise, contextual faux pas), while ensuring longer term consistency (e.g., adherence to Gricean principles). In the next section, we illustrate how interlocutors mutually develop shared conceptual spaces, systems that adapt at the same scale as the goals and needs of the interaction and that afford a certain tolerance for local ambiguity and misunderstanding.

Joint epistemic engineering

An explanatory framework of human communication needs to clarify how inherently asymmetric individuals rapidly arrive at mutually understood mental constructs from ambiguous signals that are referentially contingent on the current communicative circumstances. Existing frameworks tend to minimize those fundamental communicative obstacles, e.g., signal ambiguity (Hasson et al., 2012; Rizzolatti & Arbib, 1998), interpersonal asymmetry (Pickering & Garrod, 2013), or typological inadequacy (Gibson et al., 2019). We have recently argued, supported by neural evidence, that people overcome those obstacles because they solve, at each and every dialogue turn, a conceptual alignment challenge (Stolk et al., 2016). This alignment challenge entails that interlocutors continuously evaluate their individual conceptual structures for consistency with what is implied by the other in light of the ongoing interaction, like a reader needs to find and maintain consistent relationships between concepts conveyed in this contribution (e.g., a link between Peirce, /bæŋk/, and asymmetry). Here we zoom in on this epistemic engineering process, illustrating how interlocutors construct multi-layered communicative signals to probe and give shape to their continually changing conceptual spaces. Using two conversational settings, one verbal and one non-verbal, we show that a signal can operate across different levels: to offer cues for aligning the space of possible interpretations across interlocutors (a Peirce-signal), while asking to be recognized as communicative (a Grice-signal), and reducing uncertainty over the identity of its referent (a Shannon-signal, if the Peirce- and Grice-signals evoke the intended interpretational context). Furthermore, we show how those multi-layered signals are not processed in serial isolation. The contextual constraints suggested by the production of each Grice- and Peirce-signal accumulate and change as the interaction proceeds. By way of analogy, each signal production can be thought of as a stroboscopic flash, designed by the communicator to transiently illuminate her own rapidly shifting conceptual landscape for the benefit of the interlocutor. Across multiple turn-taking cycles, the interlocutors flash those strobes to offer a glimpse of each other's conceptual landscapes, so that they can be meshed together. In keeping with the analogy, the spectral qualities of the stroboscopic light change over time, resolving finer details of the interlocutors' contextual framework as the interaction progresses successfully, or zooming back to a coarse granularity if a new dialogue is started. The integration window over multiple signals can also change as a function of interaction status, as when we open a dialogue confidently assuming mutual knowledge of an earlier conversation with the same interlocutor, or tentatively explore a newcomer with a pleasantry. Joint epistemic engineering is meant to emphasize that interlocutors are focused on controlling the shared conceptual landscape, rather than the consistency of signals or signal-referent mappings. By embedding signals in the shared conceptual space generated through the engineering process, interlocutors rapidly zoom in on relevant features of those signals, coordinate current reference points, and prepare the ground for anchoring new reference points.

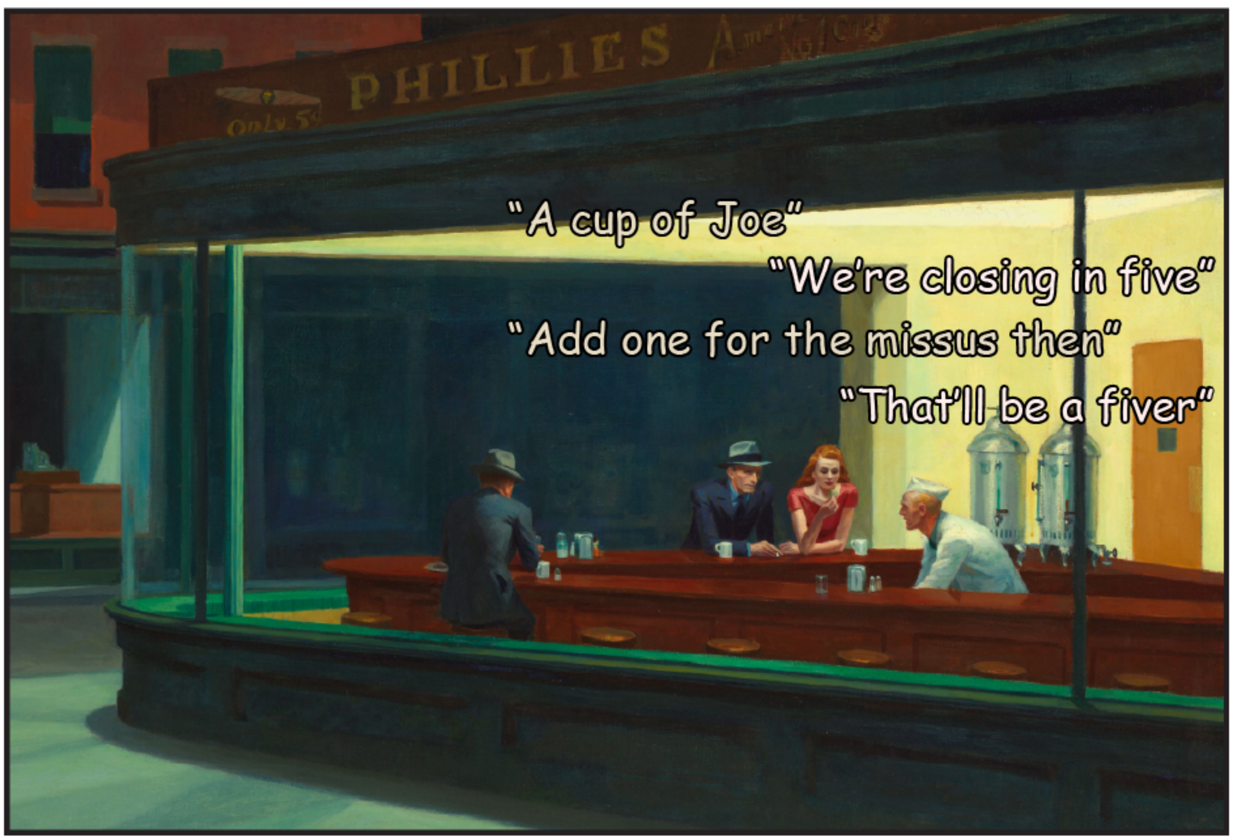

Figure 1. A prototypical human interaction, semantically-incoherent but semiotically-fluent. Artwork courtesy of the Art Institute of Chicago.

Engineering a bar conversation

It is easy to overlook how joint epistemic engineering underpins everyday interaction. Consider the bar conversation shown in Figure 1. The flow and apparent straightforward understanding between customer and bartender might mask the interpretational constraints necessary to make sense of a dialogue laced with semantic incoherence and ambiguity. Even before voicing “A cup of Joe”, the customer would need to build a scaffold of conceptual information that approximates the scaffold presumably used by the bartender to interpret his utterance (black graph structures in Figure 2A). The customer needs to select background knowledge that coffee can be ordered from bars; that bartenders may be familiar with slang words for coffee; and so on (blue background structure in Figure 2A). In fact, several possible-world scenarios might need to be prepared: the presumed bartender is a cleaner; the bar is already closed; etcetera. Accordingly, besides placing an order (a Shannon-signal), the customer’s opening statement doubles as a tacit request to probe the conceptual scaffold shared with his interlocutor (a Peirce-signal), and to be recognized as such (a Grice-signal). The bartender’s reply, “We’re closing in five”, despite bearing no semantic resemblance to the customer’s opening utterance, suffices to address the customer’s multi-layered request. He confirms his ability and willingness to process the order (addition of coffee to the shared conceptual space, Figure 2B), as well as to engage in the joint engineering process. Besides conveying recognition of communicative intent (a Grice-signal) and additional details about the bar (a Shannon-signal), the bartender’s disclosure of the approaching closing time also operates as a tacit invitation to negotiate the customer’s current request (e.g., “I’m not serving”, or “I’m serving but drink fast”) or make another (e.g., a drink for his female companion), i.e. a Peirce-signal. This multiplicity of functions illustrates how interlocutors use signals to simultaneously coordinate current and anchor new reference points in their conceptual frame, effectively coordinating a situated source of interpretational constraints for each other’s intrinsically ambiguous signals.

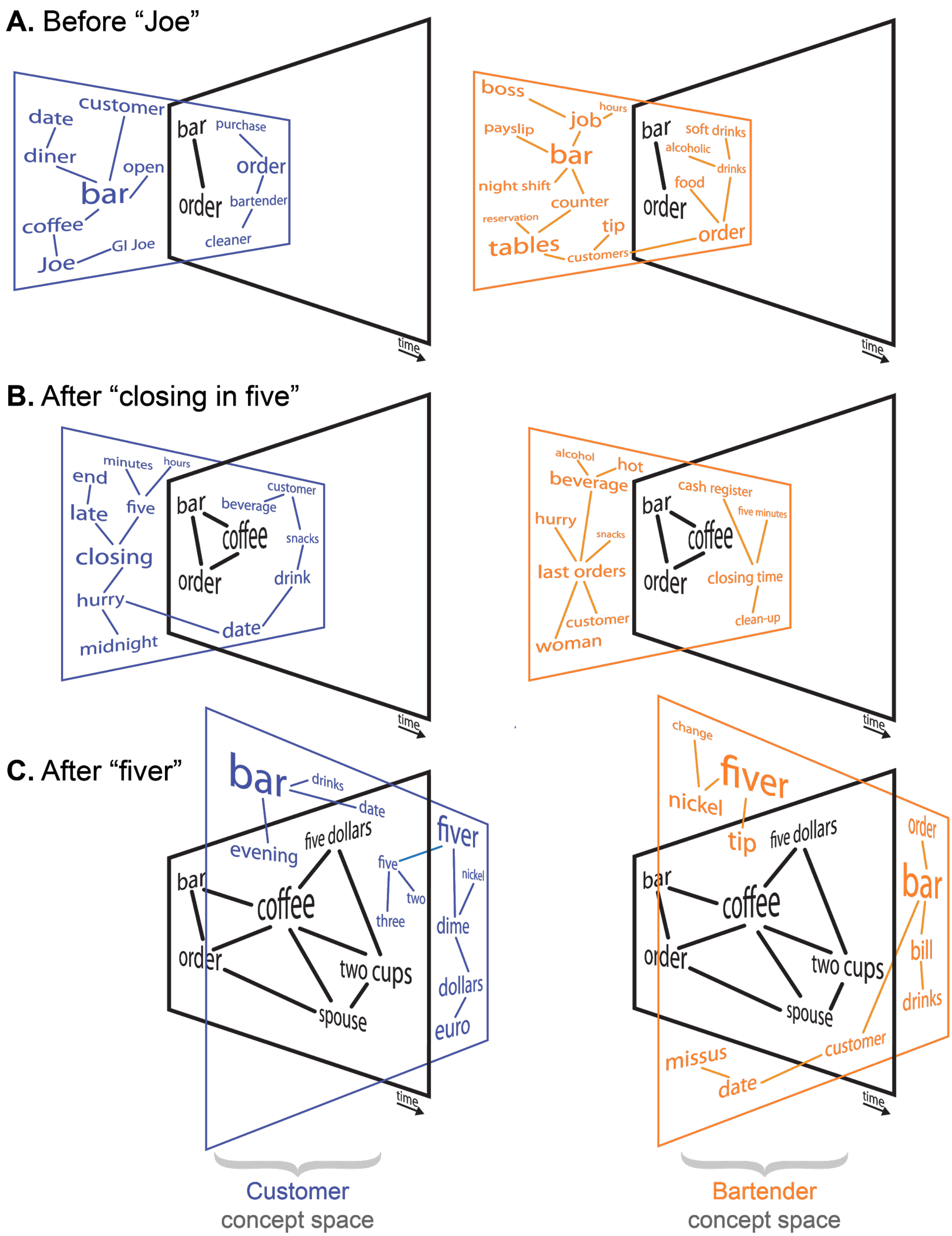

Figure 2. Epistemic dissections of three moments in the bar conversation shown in Figure 1. Collectively, the dissections illustrate how dialogue requires people to continually build a scaffold of conceptual information that

approximates the scaffold used by the other (forming a shared conceptual space, black graph structures). This joint engineering process kicks in even before the first signal is produced, and draws on individual background knowledge where possible (blue/orange graph structures). For instance, before voicing “A cup of Joe”, the customer needs to select background knowledge that coffee can be ordered from bars; that bartenders might be familiar with slang words for coffee; and so on (blue background structure, A). At the same time, several possible-world scenarios might need to be prepared: the presumed bartender is a cleaner; the bar is already closed; etcetera. Though semantically incoherent with the customer’s opening statement, the bartender’s reply, “We’re closing in five”, suffices to address the customer’s request that he can and is willing to process the order (addition of coffee to the shared conceptual space, B). Furthermore, his reply doubles as a tacit invitation to anchor new reference points in their conceptual space, e.g., a last order, details about the bar’s opening hours. As interaction unfolds, the interlocutors construct their multi-layered signals to probe and give shape to their expanding conceptual spaces, keeping those spaces aligned with one another to provide a situated source of interpretational constraints for each other’s intrinsically ambiguous signals (black conceptual structures, C).

This bar conversation also illustrates how joint epistemic engineering offers the possibility to rapidly build on a potentially large body of background knowledge (blue and orange graph structures in Figure 2). By considering background knowledge in light of the current conceptual frame, interlocutors can reference a wealth of presumably shared or readily shareable semantic content for integration into their conceptual frame. For instance, the customer’s specification of “the missus” in his reply, “Add one for the missus then”, offers the bartender with an unsolicited set of cues to the identity of his female companion, ground he could have expanded on in a subsequent exchange. Perhaps the innovation within the bartender’s response, “That’ll be a fiver”, hints at an alignment on the level of nonchalance still, consistent with the customer’s slang use. But on its own, the bartender’s response references a separate pocket of background knowledge (concept of payment, orange graph structure in Figure 2C), knowledge he expects his interlocutor to also consider when hearing his response in the context of their current conceptual frame. Needless to say, that reference would have been out of place in other moments of the exchange (e.g., before the customer’s order). These examples illustrate a means by which interlocutors can rapidly contextualize background knowledge relevant to the current situation, amalgamating potentially unrelated conceptual structures into a space of meanings and relationships between those structures. By exploiting the conceptual space they have jointly assembled and streamlined over their four utterances (black graph structures in Figure 2C), the interlocutors can rapidly generate hypotheses about novel signals, as when a customer would hear “fiver” for the first time. That epistemic inference would widely miss its target when working from the sensory data using context-free statistical regularities abstracted from speech/text corpora, as AI-based language processors do (Marcus & Davis, 2019). Lacking a cognitive architecture for joint epistemic engineering, those artificial agents would fail to interpret “fiver” in the context of the customer’s opening statement, instead limiting search for relevant context to semantically, temporally, or arithmetically closer signals in the turn-taking sequence (e.g., “in five”, “Add one”).

Engineering a communication system

It could be argued that the previous example, while plausible, remains an armchair argument with no empirical evidence. In fact, even if the utterances had been recorded from a real diner, the conversational commentary would be purely speculative and rather descriptive. Fortunately, some researchers have managed to gather experimentally controlled observations on human communication, without losing sight of its core interactive features. The rationale of the approach is that, when shared assumptions on what counts as a context and a signal are experimentally removed, then the production and interpretation of Grice- and Peirce-signals will become more apparent, revealing the mechanisms of joint epistemic engineering in the process. Although the approach might seem to rely on experimentally contrived situations (see Figure 3 for an example, discussed in more detail below), similar challenges are common when communication cannot count on a shared idiom. Take the situation experienced by the Hawaiian people when visited by James Cook in 1779. As in many other “first encounters” between different groups (Leahy, 1936), people found ways to communicate reliably and the British sailors spent a peaceful month on the island. Yet, shortly afterwards, James Cook lost his life in a confrontation, apparently a consequence of miscommunications arising from a change of context, obvious to the Hawaiian people (the shift from the season of peace to the season of war) but completely opaque to the Western visitors. It is hardly surprising that during that episode, one of the crew members became sensitive to the fixed-code fallacy in Signs and Words (see opening quote).

Several groups have recently started to re-create first encounters in the lab (without most of the risks) as a way to gain experimental access to core generative processes supporting human communicative abilities (de Ruiter et al., 2010; Galantucci, 2005; Scott-Phillips et al., 2009; Selten & Warglien, 2007). This research program, under the general banner of "experimental semiotics" (Galantucci & Garrod, 2011), has created various communication games where human participants interact using signals lacking a pre-existing conventional use (e.g., artificial words, scribbly drawings). Importantly, while signal and context accessibility are experimentally controlled, great care is usually taken to preserve the minimal condition for studying human communication, i.e. the presence of two or more people in interaction. Accordingly, jointly establishing signal meaning in those settings relies on converging on a common ground of knowledge and beliefs across interlocutors, more so than deciding the meaning of already known words and gestures. In our own work, we have frequently used the "Tacit communication game", i.e. live open-ended communicative interactions during which two players converse on a 3-by-3 digital grid using the movements of a geometric shape (de Ruiter et al., 2010). In each round of the game, the two players have to jointly reproduce a target configuration of two shapes, one for each player (a circle, rectangle, or triangle). Crucially, however, the target configuration is shown to one of the players only, the sender, see Figure 3. Given that the other player, the receiver, cannot see this configuration, the sender needs to communicate the receiver's target position to the receiver while also ensuring that the final location and orientation of their own shape is as specified by the target configuration. The only means available to the sender for communicating with the receiver is by moving the sender's own shape around the grid using horizontal and vertical translations, and 90° clockwise rotations controlled by button presses on a handheld game controller. The only means available to the receiver for completing the configuration is by inferring the target location and orientation of his or her own shape on the basis of the movements of the sender, and positioning it accordingly using a second handheld game controller. As the receiver cannot ask for clarification, the only indication of whether her interpretation of the sender's signals was correct is feedback on the token's position received at the end of each trial. Moreover, the sender not only has to monitor whether the receiver got to the target location, but also whether she is adhering to the tacitly agreed-upon communicative strategy or making any changes in the way communicative signals have been used so far.

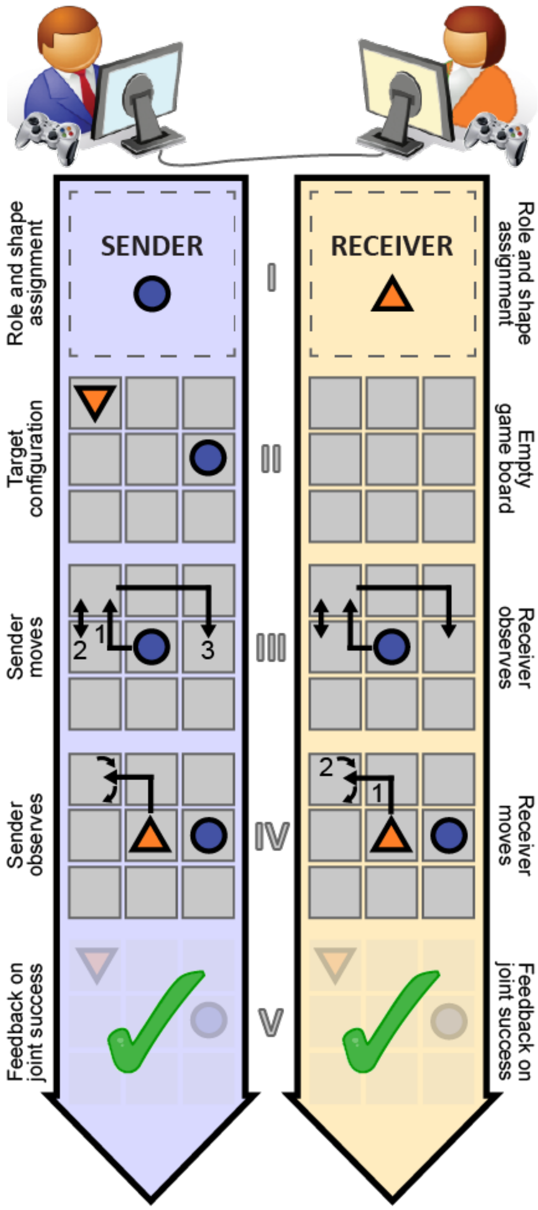

Figure 3. Tacit communication game. The joint goal of sender (left) and receiver (right) is to reproduce a target configuration of their two given shapes on a digital game board. This target configuration is shown to the sender only (event II). Given that the receiver cannot see the target configuration, a successful interaction requires the sender and receiver to construct and comprehend how the behavior of the sender's shape indicates the location and orientation of the receiver's shape (a 'communicative signal', event III). Feedback in the form of a green check mark or red cross is presented to sender and receiver to indicate whether or not they successfully reproduced the target configuration (communicative success, event V). A benefit of the novel communicative setting is that the shape behaviors lack conventional meaning, providing reliable access to the core interpersonal process of how two humans converge on a shared meaning. An online version of the game can be played at www.MutualUnderstanding.nl.

There are no a priori correct communication strategies in the Tacit communication game nor can the receiver solve the task by simply reproducing the movements of the sender's shape. Analogous to everyday multimodal communication involving ambiguous signals, the receiver needs to rapidly zoom in on relevant features of the sender's movements and infer a possible referent (a shape configuration) on the basis of those features. Example features include brief movement "pauses" on the receiver's target location, distinguishing it from other locations visited on the grid, as well as rotational movements, directional movements, and repeated stepping in and out from the receiver's target location ("wiggling"), with the latter three movement types typically used to indicate target orientation. It is tempting to typologically assign those movement features to signals' referents as if they constitute optimal solutions to (parts of) a problem, but this would miss the local principles that govern their interpretations (Garrod & Anderson, 1987). For instance, it cannot be determined from a wiggling behavior alone whether it is intended to emphatically indicate the direction in which the receiver's shape needs to point, or whether the number of wiggles corresponds to the number of clockwise rotations the receiver needs to make to reach the target orientation, and some pairs use both of these strategies during a single game. Furthermore, interlocutors dynamically converge on solutions that may not reliably generalize to other contexts. A case in point is provided by a pair still getting acquainted with the task instructions. Not yet realizing the need to communicate the receiver's target location, the sender of this particular pair moved straight to the sender's own target location, thus requiring the receiver to make a blind guess regarding the receiver's target location. Against the odds, the receiver chose the correct location diagonally opposite to the sender's final position. Having apparently become mutually aware of this accidental agreement, the pair kept employing their feature-less communicative strategy on similar problems, even when they had developed alternative solutions with more generalization power. These observations indicate how in this game, as in the bar conversation, the trajectory from a signal to its referent is contingent on the communicative context people develop together. Indeed, the same communicative behavior can be used by different pairs to convey different meanings and, as we will observe next, can have different meanings in different interactions of the same pair.

To illustrate how joint epistemic engineering could support rapidly understood new signal-referent trajectories, even when a different trajectory exists for a signal, we here turn to the case of the “missing wiggle”. Before doing so, it should be mentioned that to drive interlocutors to continuously develop and coordinate their shared conceptual spaces, the Tacit communication game increases the task difficulty across successive interactions. This is achieved by introducing deliberate mismatches between the geometric characteristics of the pairs’ shapes, including situations where a receiver’s target orientation is incompatible with the pair’s current strategy. For instance, if a sender was able to successfully communicate the target orientation of the receiver’s shape by rotating their own shape, this strategy would be negated when the sender’s shape changed to a circle and the receiver’s to a triangle. The sender would then have to find a new way to indicate the target orientation of the receiver’s shape, because rotations of a circle are not visible. A further level of difficulty is introduced by maintaining the same circle-triangle shape combination, but having the receiver’s triangle point outside the grid. If the sender had previously used a signal that involved repeatedly stepping in and out of the receiver’s target location in the direction of the target orientation (“wiggling”, number 2 double-headed arrow in Figure 3), an interaction where the receiver’s triangle pointed outside of the grid would be impossible to solve using that signal. Returning to our case, several pairs started to use a “pause” strategy to indicate that the triangle needed to point outside of the grid. Those pairs had previously converged on using movement pauses on receiver target locations to selectively indicate the communicative relevance of those grid locations, thus with no obvious role for pauses in conveying target orientation other than the default orientation. Hence, through a lens of associative learning, the use of a pause strategy would unambiguously mean that a triangle needed to be positioned in its default upward position, rather than outward facing. How then could pairs have rapidly converged on a new use for an existing signal in the absence of any failures? A closer look at a pair’s recent communicative history hints at a conceptual shift as to what movement features the pair mutually considered relevant in their communication. Specifically, the pair had recently converged on using a wiggle strategy to indicate the direction in which a triangle needed to point (as in Figure 3), applying this strategy even in cases where a triangle needed to point upward and previously a pause would have sufficed. In doing so, wiggling had not only become the prevalent feature for indicating target orientation in this pair, but the conceptual shift in feature relevance had implicitly relaxed the relationship between pauses and default orientations. Consequently, pauses could be introduced by the sender to signify orientation in situations where he could safely assume that the receiver would infer wiggles could not (i.e. a missing wiggle, see dashed lines in Figure 4), such as with outward-facing triangles. This epistemic inference illustrates how a shift in conceptual space can cause a signal to acquire a different meaning in (apparently) a single instance, as when a pause becomes a missing wiggle.

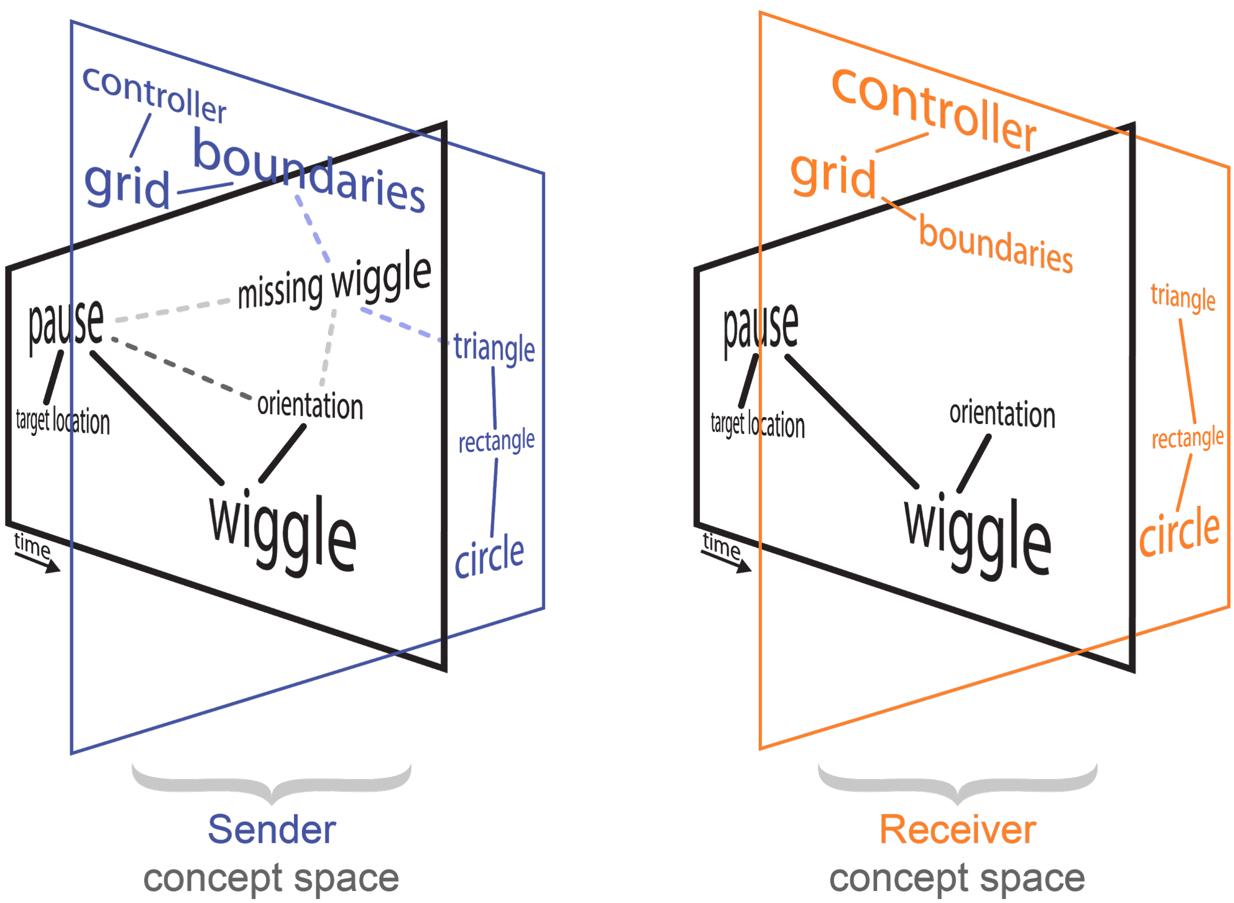

Figure 4. Epistemic dissection capturing the moment a signal acquires a new meaning in the Tacit communication game. The figure on the left captures the moment the sender used a pause, previously used exclusively to signal target location (solid black line), to additionally signal orientation (dashed grey line), a function previously reserved for wiggles (solid black line). To understand this new use for behaviorally identical pauses (pause as a “missing wiggle”), the receiver would need to infer the implied change in their conceptual space driving that proposal. For instance, he would need to infer that the sender faced a situation in which wiggles could no longer be used to signal orientation, such as with an outward-facing triangle at the grid’s boundary. He would also need to consider conceptual shifts supporting that new wiggle-less orientation-signaling strategy (dashed lines radiating out from

"missing wiggle", left figure). If he then manages to convey agreement about that epistemic inference to the sender, the players can strengthen or add connections in their conceptual space supporting the sender's proposal (dashed grey lines on the left becoming solid in both left and right figures).

What does the brain care about?

The previous sections provide prima facie evidence to move beyond views of human communication as information transfer. Humans do not share signals or a fixed-code for encoding and decoding thoughts from those signals. Nor can they realistically be expected to disambiguate communicative signals based on fixed contextual cues or simple clarification procedures. Rather, humans might share a capacity to jointly build and coordinate a conceptual frame of reference informed by what they commonly presume is known and believed by the other under the current circumstances. By selecting and interpreting signals in the context of the current conceptual frame, signals have a fairer chance at being understood in the way they are intended at the moment of delivery, whether during ordinary conversation or during "first encounters" deprived of a common language (cf. Figures 2 & 4). In this light, and by extension of Peirce, signals can take any form and are merely a means to probe and bias the conceptual frame shared with the other. As interaction unfolds, interlocutors continuously update their conceptual frames, keeping their mutually inferred thoughts aligned with the current situation and with one another. Here we discuss recent work that has begun to empirically interrogate these theoretical issues at the implementation level by combining neuroscientific techniques with controlled studies of live social interaction, i.e. the Tacit communication game outlined above. To date, there are three major insights.

First, both production and comprehension of mutually understood communicative signals are supported by a right-lateralized frontotemporal network, shown in Figure 5A (Stolk et al., 2013). This network has been shown to contribute to pragmatic and mental state inferences when communicating with linguistic means, suggesting a role for right frontotemporal cortex in embedding utterances in their conversational context (Bašnáková et al., 2014; Beeman, 1993; Sabbagh, 1999). Second, neural activity in this right-lateralized network predates in time the production and comprehension of communicative signals, i.e. before the initial dashed lines in Figure 5A (Stolk et al., 2013; Stolk, Noordzij, Verhagen, et al., 2014). This observation fits with the notion that communicative behaviors are selected and interpreted within the conceptual frame of reference people in dialogue develop together. Interference studies of anterior frontal and posterior temporal nodes of the network confirm and qualify the contribution of those cortical regions. Ventromedial prefrontal lesion patients, known for their social conduct deficits (Beer et al., 2006), were found lacking the specific ability to constrain their ongoing communication with knowledge implied by the current communicative circumstances, including the presumed abilities of an interlocutor (Stolk et al., 2015). Transient disruption of neural function in the right posterior superior temporal sulcus diminishes the ability to constrain the interpretation of communicative behaviors using knowledge abstracted from the recent communicative history (Stolk, Noordzij, Volman, et al., 2014). These causal observations emphasize the constraining role of contingently updated conceptual knowledge in social interaction.

Third, neural activity in the same right-lateralized frontotemporal network becomes synchronized across interlocutors due to the accumulation of shared conceptual knowledge at a temporal scale independent from individual communicative signals (Stolk, Noordzij, Verhagen, et al., 2014). This finding is a merit of using novel and controlled interactive settings, offering the possibility to experimentally manipulate and track interlocutors' shared signal-referent mappings across interactions. For instance, interlocutors understood one another better when they encountered situations for which they had previously converged on a shared communicative solution, reflected in greater communicative success and neural activity in the right-lateralized frontotemporal network. One subregion of the network, the superior temporal gyrus (STG), exhibited neural dynamics matched to the behavioral dynamics of communicative success over the course of the experiment. Interestingly, those neural dynamics were synchronized across interlocutors, but only over episodes when interlocutors developed shared communicative solutions, and at a temporal scale independent of the occurrence of the communicative behaviors themselves (coherent neural dynamics in Figure 5B). Conversely, this neural synchronization was reduced over episodes when interlocutors could rely on previously shared communicative solutions, thus, communicative episodes that did not require adjusting knowledge. These observations define a neural basis for the notion that we continuously update and sharpen our conceptual knowledge according to the recent history of the communicative interaction. Moreover, the simultaneous neural changes in superior temporal gyrus during episodes of knowledge adjustment across interlocutors suggest that we mutually coordinate those adjustments.

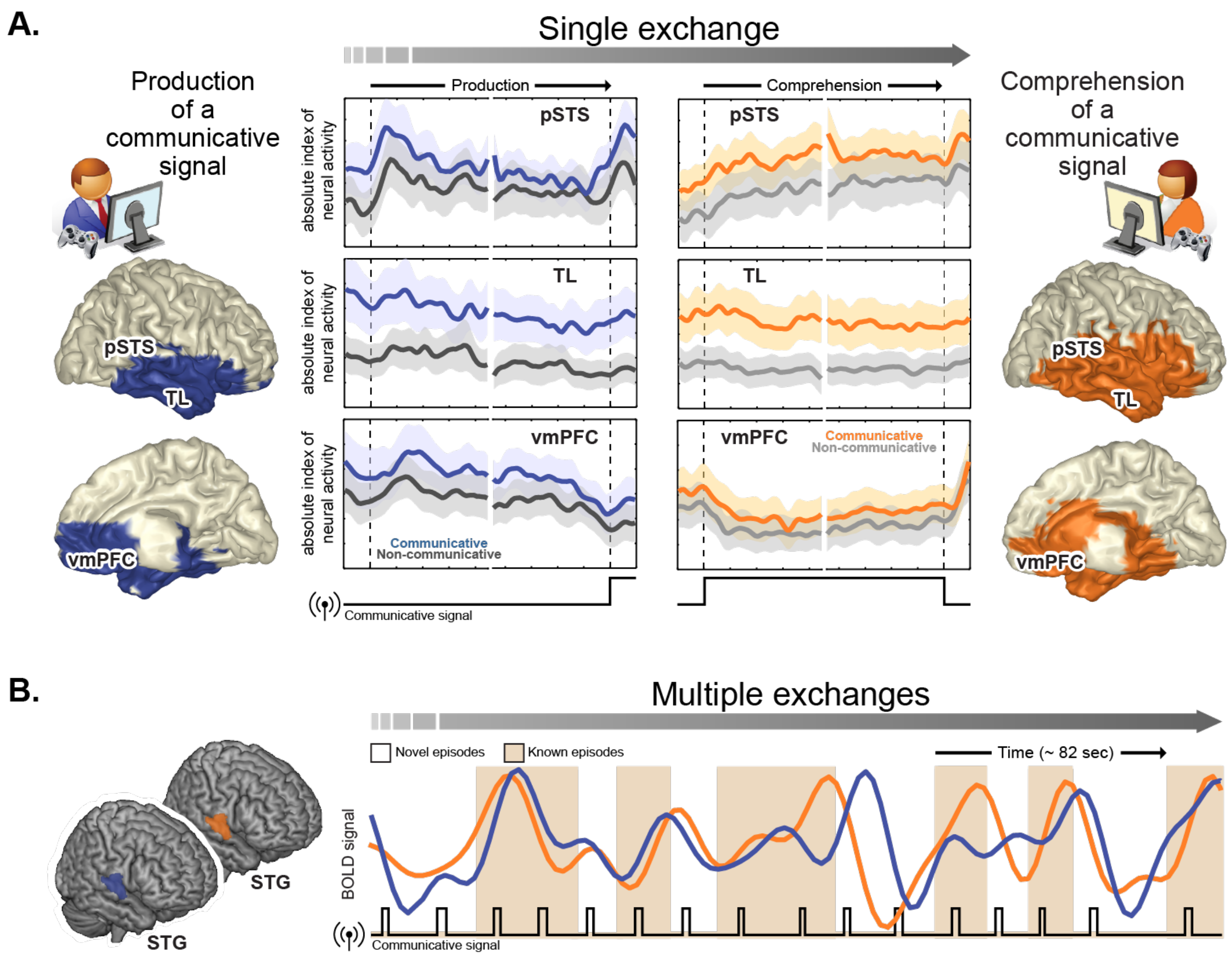

Figure 5. Neural dynamics of sharing conceptual spaces within a single communicative exchange (A) and over multiple exchanges (B). Brain regions of a right-lateralized frontotemporal network supporting both production and comprehension of communicative signals (blue and orange, respectively) are already upregulated before a communicative signal is produced or comprehended (before the initial dashed in lines in A). Activity in one subregion, the superior temporal gyrus (STG, shown in B), tracked communicative success over the course of the experiment (not shown), and was coherent across interlocutors at a temporal scale independent from individual communicative signals (blue and orange traces). Adapted from (Stolk et al., 2016).

The picture that emerges from these brain observations is one in which the meaning of a communicative signal is not a property of the signal, nor of individuals interpreting that signal. Rather, meaning is a property of conceptual knowledge inferred from prior beliefs and interactive behaviors embedded in the recent communicative history, and kept aligned with the unfolding interaction to constrain communication. What neurophysiological mechanism allows a sustained yet adjustable influence of conceptual knowledge on transient signal production and comprehension? A clue is provided by the right-hemispheric frontotemporal structures found to support communication. Densely spaced neuronal populations with overlapping dendritic fields and thin axonal myelination endow those frontotemporal association areas with a relatively large-world architecture compared to their left hemisphere homologues (Chance, 2014; Hutsler & Galuske, 2003; Jung-Beeman, 2005). Highly interconnected populations of neurons might confer conceptual processing at varying levels of abstraction, offering a neuroanatomical infrastructure to flexibly scaffold and search for relevant context in communication. Another clue is provided by the spectrotemporal nature of neural activity observed over the frontotemporal cortex when people interact. Temporally, frontotemporal regions are tonically upregulated during communicative interactions, showing strikingly similar phasic neural dynamics during the production and comprehension of both communicative and non-communicative events (neural dynamics of Figure 5A; for details, see (Stolk et al., 2013)). This observation is consistent with an increasing body of evidence showing that contextual demands can modulate ongoing neural activity yet retain responsiveness to event-related neural processing (Abitbol et al., 2015; Musall et al., 2019). Indeed, the tonic upregulation had measurable behavioral consequences on communicative performance (Stolk et al., 2013). Spectrally, the neural upregulation over the frontotemporal cortex had an extremely broad profile (not shown in Figure 5). Broadband shifts of neural activity are thought to reflect changes in the mean firing rates of neuronal populations, and their antecedence to observable events might be instrumental to integrate driving afferences with contextual information. Specifically, ongoing contextual inputs can temporarily hold selective neurons in an excitable state, increasing the probability of those neurons propagating event-related inputs, effectively integrating information associated with the input streams (Aru et al., 2020). This neuronal mechanism, based on upregulated broadband neural activity in the right frontotemporal cortex, might provide a neural marker of the conceptual frame of reference people build together during social interaction (Stolk et al., 2013).

What could possibly go wrong?

In this contribution, we have provided theoretical and empirical arguments for why the field can and should move beyond views of human communication as information transfer. There are also clinically salient reasons for overcoming those pre-Darwinian views, based on how communication breaks down in the presence of a relatively intact command of language. For instance, individuals with damage to the right hemisphere may respond appropriately to clearly structured questions requiring specific answers, e.g., "Where do you live?". Yet they produce irrelevant and verbose answers (in a Gricean sense) to open-ended questions such as "What did you do yesterday?". Patients seem unable to assemble their utterances into a coherent narrative, instead wending their way through a maze of disassociated detail, without answering the question (Myers, 1979). Right hemisphere communicative deficits are known to permeate language processing at multiple levels, from difficulties with identifying the main theme in a conversation to using prosodic speech modulations (Ross & Monnot, 2008). Comparable deficits have been noted in patients suffering frontotemporal dementia, a neurodegenerative disorder marked by deterioration of frontotemporal cortical structures. Analyses of the neuropsychology of frontotemporal dementia patients reveal that coherent conversations as well as intrinsically motivated social relationships are affected when the frontal lobe and right temporal pole degenerate (Mates et al., 2013). These clinical observations converge with the above-mentioned empirical findings in linking a right-lateralized frontotemporal network to the processing of conceptual knowledge of the communicative context, additionally highlighting a role for this processing in sustaining social bonds. Precise quantification of the communicative alterations in patients with frontotemporal dementia may prove a fruitful research direction. By exploiting the natural variability in location and extent of neurodegeneration in this pathology, it might become possible to triangulate the relation between brain tissue loss and deficits in creating a coherent representation of the conversational context (i.e. the black graph structures of Figures 2 & 4).

Ultimately, it will be important to make systematic comparisons between patient groups with communicative alterations. This pathophysiological comparative approach might isolate cognitive phenotypes that unify apparently disparate disorders (e.g., schizophrenia, autism spectrum disorder, frontotemporal dementia). For instance, it might emerge that patients with either of those disorders have problems in creating and maintaining a shared conceptual space with an interlocutor. Recent evidence from our group suggests that this alteration is present in cognitively able adults on the autism spectrum (Wadge et al., 2019). Despite otherwise indistinguishable performance from neurotypical pairs, pairs of individuals diagnosed with Autism Spectrum Disorder (ASD) had lower communicative success in the Tacit communication game. This communicative impairment was not simply a consequence of reduced cognitive flexibility or social motivation, as individuals with autism showed a similar propensity to change their communicative behaviors following a misunderstanding with their communication partner. They even spontaneously placed more emphasis on communicatively relevant portions of their behavior for the benefit of their partners, as neurotypical partners would do. Crucially, however, individuals with autism struggled to rapidly converge on a shared conceptualization of their communicative behaviors with their communication partner. This communicative misalignment predicted communicative impairment across all pairs and was greatest when the conceptualizations depended on the unique communicative context established through past interaction with their partners. These findings provide causal evidence for the notion that the meaning of a communicative signal is best understood within the conceptual frame of reference people in dialogue develop together. As epitomized by the largely segregated interaction spaces between individuals with autism (Figure 6, cf. proportion of single-color clusters in the pairwise interaction trajectories), neurotypical individuals navigate and constrain the vast space of meaning by continuously considering and aligning to recent signals from their partner (in a Peircian sense). This communicative misalignment explains why autistic individuals are vulnerable in everyday social situations entailing fleeting and interaction-specific ambiguities, but succeed in social cognition tests where they can capitalize on stereotyped contextual cues to resolve ambiguities (Branigan et al., 2016; Hahn et al., 2015). Precise characterization of how individuals with and without autism differentially integrate semantically distant knowledge structures into a coherent conceptual space might provide a new window into understanding autistic communication, and how signals derive their meaning from the communicative context in which they are embedded.

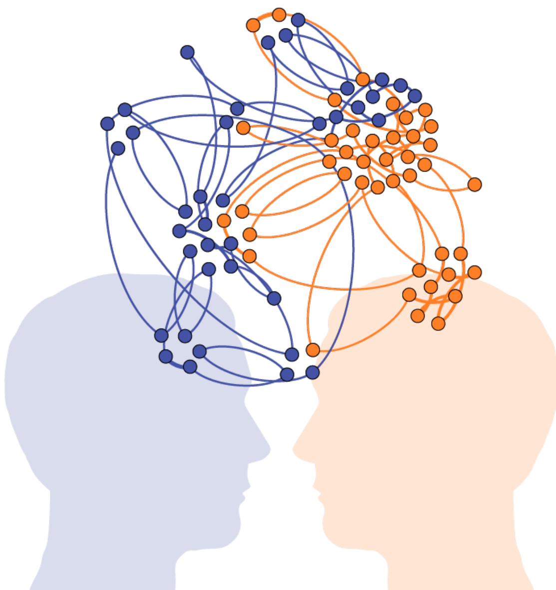

Figure 6. Network visualizations of pairwise interaction trajectories by a pair of individuals on the autism spectrum (left) and a neurotypical pair (right). The nodes represent communicative signals constructed by two individuals during a series of interactions in the Tacit communication game. The nodes are clustered to show signals that were used repeatedly. The colored edges connect each individual's consecutively produced signals over the course of the interaction (blue for one individual, orange for the other). It can be seen that individuals with autism showed more individual exploration of their interaction space, as indicated by relatively large clusters of individually used signals (clusters of nodes that were connected only by edges of a single color) and small clusters of jointly used signals. In contrast, the neurotypical pair navigated the interaction space by continuously considering and aligning to recent signals from their partner, embedding those signals in a strongly interconnected space of meaning and relationships between one another's signals. ASD, Autism Spectrum Disorder. Adapted from (Wadge et al., 2019).

The theoretical and empirical observations presented in this chapter might be seen as exceptional situations that do not generalize to the daily interactions of a linguistic community. On the contrary, the same mechanisms bootstrapping communication in experimental circumstances must be at work when people learn to use conventional communicative signals in their daily interactions. Although the role of the dynamic communicative context in language learning has been largely understudied, several pieces of evidence point in this direction. For instance, while language processing is largely left-lateralized throughout life, necessary contributions from the right hemisphere may be seen early in life, gradually decreasing through childhood (Bates et al., 2001; Olulade et al., 2020). Moreover, the right hemisphere may support basic language skills when the left hemisphere is compromised (Newport et al., 2017). To date, there is no evidence for the opposite, with the left hemisphere seemingly not able to compensate for lost conversational abilities after early right hemisphere damage (Save-Pédebos et al., 2016). As noted elsewhere (Sabbagh, 1999), this pattern of recovery is reminiscent of the pattern of linguistic and communicative development in autism. Vocabulary development is delayed but occurring in autistic children, yet communicative deficits persist through adulthood (Tager-Flusberg & Anderson, 1991; Wadge et al., 2019), as they would in right hemisphere patients. We infer from its comparably small-world architecture that the left hemisphere is less able to cope with the flexibility demands of everyday interaction. Widely spaced neuron populations with little overlapping dendritic fields and greater levels of myelination endow the left hemisphere with functional stability and efficiency in terms of perturbation-resistant network hubs and shorter path lengths. Accordingly, the left hemisphere might be better equipped to generalize recurrent patterns of association (e.g., semantic knowledge), when scaffolded by the conceptually flexible right hemisphere. These stereotyped dependencies, on their turn, arguably help communication by providing a predictable shared structure and cues to a meaning (blue and orange background structures of Figure 2).

Conclusion

Human communication is often explained in terms of sending and receiving signals, presupposing that interlocutors share an optimal and stable code to insert and extract meaning from those signals. This pre-Darwinian intuition has percolated into modern frameworks of language and communication, where it remains largely indifferent to fundamental communicative obstacles facing signal use and interpretation in everyday dialogue, including signal ambiguity, interpersonal asymmetry, and typological inadequacy. We have argued that people in dialogue find ways around those obstacles by solving a different control problem than the one postulated in coding-decoding frameworks. Rather than working from the sensory data using context-free statistics of those data, people jointly control a conceptual frame of reference through the production of multi-layered signals. Besides reducing uncertainty over the identity of a referent in the interlocutors' minds (a Shannon-signal), a multi-layered signal asks interlocutors to manage that declaration of communicative intent (a Grice-signal) and to coordinate the space of possible interpretations across interlocutors (a Peirce-signal). The inferential process underlying multi-layered signals is contingent on the interaction history and provides the basis for joint epistemic engineering - the act of building a shared context that makes Shannon-signals similarly interpretable through the generation of Peirce- and Grice-signals. Recent empirical findings from our group, obtained in experimental settings that respect the minimal conditions for studying communication, suggest that the brain solves those fundamental communicative obstacles by continuously tracking interlocutors' conceptual landscapes, so that they can be meshed together. This shared conceptual space provides critical context for using and interpreting multi-layered signals, and is jointly coordinated and updated during social interaction through a tonic upregulation of neural populations in right-hemispheric frontotemporal areas. Furthermore, that coordinated updating is specifically altered in interlocutors on the autism spectrum. These considerations open the way to precisely characterize neural mechanisms that can implement the functions postulated by the joint epistemic engineering framework. For instance, it remains mechanistically and computationally unclear how a signal can modify the context that makes its interpretation communicatively appropriate. It is also unclear how to precisely control the various degrees of freedom implicit in the suggestions that interlocutors continuously adjust the integration window and conceptual granularity of their contextual representations. Future work should aim to understand how the elusive conceptual space underlying human communication is generated in our brains, how its neural implementation is altered in disorders of human communication, as well as how joint epistemic engineering can bootstrap communication when infants learn to use linguistic signals in their daily interactions.

References

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B. J., Bouret, S., & Pessiglione, M. (2015). Neural mechanisms underlying contextual dependency of subjective values: converging evidence from monkeys and humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(5), 2308–2320.

Aru, J., Suzuki, M., & Larkum, M. E. (2020). Cellular Mechanisms of Conscious Processing. *Trends in Cognitive Sciences*, *24*(10), 814–825.

Bang, D., Ershadmanesh, S., Nili, H., & Fleming, S. M. (2020). Private–public mappings in human prefrontal cortex. In *eLife* (Vol. 9). https://doi.org/10.7554/elife.56477

Bašnáková, J., Weber, K., Petersson, K. M., van Berkum, J., & Hagoort, P. (2014). Beyond the Language Given: The Neural Correlates of Inferring Speaker Meaning. In *Cerebral Cortex* (Vol. 24, Issue 10, pp. 2572–2578). https://doi.org/10.1093/cercor/bht112

Bates, E., Reilly, J., Wulfeck, B., Dronkers, N., Opie, M., Fenson, J., Kriz, S., Jeffries, R., Miller, L., & Herbst, K. (2001). Differential Effects of Unilateral Lesions on Language Production in Children and Adults. In *Brain and Language* (Vol. 79, Issue 2, pp. 223–265). https://doi.org/10.1006/brln.2001.2482

Beeman, M. (1993). Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. *Brain and Language*, *44*(1), 80–120.

Beer, J. S., John, O. P., Scabini, D., & Knight, R. T. (2006). Orbitofrontal cortex and social behavior: integrating self-monitoring and emotion-cognition interactions. *Journal of Cognitive Neuroscience*, *18*(6), 871–879.

Boorman, E. D., & Rushworth, M. F. S. (2009). Conceptual Representation and the Making of New Decisions. In *Neuron* (Vol. 63, Issue 6, pp. 721–723). https://doi.org/10.1016/j.neuron.2009.09.014

Branigan, H. P., Tosi, A., & Gillespie-Smith, K. (2016). Spontaneous lexical alignment in children with an autistic spectrum disorder and their typically developing peers. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*(11), 1821–1831.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(6), 1482–1493.

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: the development of children's social understanding within social interaction. *The Behavioral and Brain Sciences*, *27*(1), 79–96; discussion 96–151.

Chance, S. A. (2014). The cortical microstructural basis of lateralized cognition: a review. *Frontiers in Psychology*, *5*, 820.

Clark, H. H. (1997). Dogmas of understanding. In *Discourse Processes* (Vol. 23, Issue 3, pp. 567–598). https://doi.org/10.1080/01638539709545003

Clark, H. H. (2020). Anchoring Utterances. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12496

Clark, H. H., & Marshall, C. R. (1981). *Definite knowledge and mutual knowledge*. https://philpapers.org/archive/CLADKA.pdf

Conway, B. R. (2012). Color consilience: color through the lens of art practice, history, philosophy, and neuroscience. *Annals of the New York Academy of Sciences*, *1251*, 77–94.

Corballis, M. C. (2017). Language Evolution: A Changing Perspective. *Trends in Cognitive Sciences*, *21*(4), 229–236.

Danchin, E., Giraldeau, L.-A., Valone, T. J., & Wagner, R. H. (2004). Public information: from nosy neighbors to cultural evolution. *Science*, *305*(5683), 487–491.

Dawkins, R., & Krebs, J. R. (1978). Animal signals: information or manipulation. In J. R. Krebs & N. B. & Davies (Eds.), *Behavioural ecology: an evolutionary approach* (pp. 282–309). Oxford: Blackwell Scientific.

de Ruiter, J. P., Noordzij, M. L., Newman-Norlund, S., Newman-Norlund, R., Hagoort, P., Levinson, S. C., & Toni, I. (2010). Exploring the cognitive infrastructure of communication. In *Interaction Studies* (pp. 51–77). https://doi.org/10.1075/bct.45.04rui

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., Gisladottir, R. S., Kendrick, K. H., Levinson, S. C., Manrique, E., Rossi, G., & Enfield, N. J. (2015). Universal Principles in the Repair of Communication Problems. In *PLOS ONE* (Vol. 10, Issue 9, p. e0136100). https://doi.org/10.1371/journal.pone.0136100

Dryer, M. S. (1991). SVO languages and the OV: VO typology. In *Journal of Linguistics* (Vol. 27, Issue 2, pp. 443–482). https://doi.org/10.1017/s0022226700012743

Fillenbaum, S. (1974). Pragmatic normalization: Further results for some conjunctive and disjunctive sentences. In *Journal of Experimental Psychology* (Vol. 102, Issue 4, pp. 574–578). https://doi.org/10.1037/h0036092

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737–767.

Galantucci, B., & Garrod, S. (2011). Experimental Semiotics: A Review. In *Frontiers in Human Neuroscience* (Vol. 5). https://doi.org/10.3389/fnhum.2011.00011

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218.

Gaskell, G. M. (2007). *The Oxford Handbook of Psycholinguistics*. Oxford University Press, USA.

Gibson, E., Futrell, R., Piantadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How Efficiency Shapes Human Language. *Trends in Cognitive Sciences*, *23*(12), 1087.

Goldberg, A. E. (2019). *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton University Press.

Grice, H. P. (1957). Meaning. In *The Philosophical Review* (Vol. 66, Issue 3, p. 377). https://doi.org/10.2307/2182440

Grice, H. P. (1975). Logic and Conversation. In *Speech Acts* (pp. 41–58). https://doi.org/10.1163/9789004368811_003

Hahn, N., Snedeker, J., & Rabagliati, H. (2015). Rapid Linguistic Ambiguity Resolution in Young Children with Autism Spectrum Disorder: Eye Tracking Evidence for the Limits of Weak Central Coherence. *Autism Research: Official Journal of the International Society for Autism Research*, *8*(6), 717–726.

Harris, R. (1998). *Introduction to Integrational Linguistics*. Pergamon Press.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114–121.

Hutsler, J., & Galuske, R. A. W. (2003). Hemispheric asymmetries in cerebral cortical networks. *Trends in Neurosciences*, *26*(8), 429–435.

Jefferson, G. (1988). On the Sequential Organization of Troubles-Talk in Ordinary Conversation. In *Social Problems* (Vol. 35, Issue 4, pp. 418–441). https://doi.org/10.1525/sp.1988.35.4.03a00070

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. In *Trends in Cognitive Sciences* (Vol. 9, Issue 11, pp. 512–518). https://doi.org/10.1016/j.tics.2005.09.009

Kendrick, K. H., & Drew, P. (2016). Recruitment: Offers, Requests, and the Organization of Assistance in Interaction. In *Research on Language and Social Interaction* (Vol. 49, Issue 1, pp. 1–19). https://doi.org/10.1080/08351813.2016.1126436

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

Leahy, M. (1936). The Central Highlands of New Guinea. In *The Geographical Journal* (Vol. 87, Issue 3, p. 229). https://doi.org/10.2307/1786763

Levinson, S. (2006). On the human interactional engine. In N. Enfield & S. Levinson (Eds.), *Roots of Human Sociality* (pp. 39–69). Berg.

Lewis, D. K. (1969). *Convention: a Philosophical Study, by David K. Lewis*.

Lotto, R. B., & Purves, D. (2000). An empirical explanation of color contrast. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(23), 12834–12839.

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Vintage.

Mates, A. W., Mikesell, L., & Smith, M. S. (2013). *Language, Interaction, and Frontotemporal Dementia: Reverse Engineering the Social Mind*. Equinox.

Mayr, E. (1999). *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. Harvard University Press.

Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, *18*(10), 512–519.

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677–1686.

Myers, P. S. (1979). Profiles of communication deficits in patients with right cerebral hemisphere damage: Implications for diagnosis and treatment. *Clinical Aphasiology: Proceedings of the Conference 1979*, 38–46.

Newport, E. L., Landau, B., Seydell-Greenwald, A., Turkeltaub, P. E., Chambers, C. E., Dromerick, A. W., Carpenter, J., Berl, M. M., & Gaillard, W. D. (2017). Revisiting Lenneberg's Hypotheses About Early Developmental Plasticity: Language Organization After Left-Hemisphere Perinatal Stroke. *Biolinguistics*, *11*, 407–422.

Noveck, I. A., & Reboul, A. (2008). Experimental pragmatics: a Gricean turn in the study of language. *Trends in Cognitive Sciences*, *12*(11), 425–431.

Olulade, O. A., Seydell-Greenwald, A., Chambers, C. E., Turkeltaub, P. E., Dromerick, A. W., Berl, M. M., Gaillard, W. D., & Newport, E. L. (2020). The neural basis of language development: Changes in lateralization over age. In *Proceedings of the National Academy of Sciences* (p. 201905590). https://doi.org/10.1073/pnas.1905590117

Peirce, C. S. (1931). *Collected Papers of Charles Sanders Peirce* (Hartshorne, Charles Weiss, Paul Burks, Arthur W (ed.)). Thoemmes Continuum.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences*, *36*(4), 329–347.

Port, R. F. (2010). Language as a Social Institution: Why Phonemes and Words Do Not Live in the Brain. In *Ecological Psychology* (Vol. 22, Issue 4, pp. 304–326). https://doi.org/10.1080/10407413.2010.517122

Rendall, D., & Owren, M. J. (2013). Communication without meaning or information: abandoning language-based and informational constructs in animal communication theory. In U. E. Stegmann (Ed.), *Animal Communication Theory* (pp. 151–188).

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194.

Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain and Language*, *104*(1), 51–74.

Sabbagh, M. A. (1999). Communicative intentions and language: evidence from right-hemisphere damage and autism. *Brain and Language*, *70*(1), 29–69.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. In *Language* (Vol. 50, Issue 4, p. 696). https://doi.org/10.2307/412243

Save-Pédebos, J., Pinabiaux, C., Dorfmuller, G., Sorbets, S. F., Delalande, O., Jambaqué, I., & Bulteau, C. (2016). The development of pragmatic skills in children after hemispherotomy: Contribution from left and right hemispheres. In *Epilepsy & Behavior* (Vol. 55, pp. 139–145). https://doi.org/10.1016/j.yebeh.2015.12.013

Schelling, T. C. (1960). *The Strategy of Conflict*. Harvard University Press.

Scott-Phillips, T. C. (2017). Pragmatics and the aims of language evolution. In *Psychonomic Bulletin & Review* (Vol. 24, Issue 1, pp. 186–189). https://doi.org/10.3758/s13423-016-1061-2

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*(2), 226–233.

Selten, R., & Warglien, M. (2007). The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7361–7366.

Shannon, C. E. (1948). A Mathematical Theory of Communication. In *Bell System Technical Journal* (Vol. 27, Issue 3, pp. 379–423). https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Blackwell Publishers.

Stivers, T. (2010). An overview of the question–response system in American English conversation. In *Journal of Pragmatics* (Vol. 42, Issue 10, pp. 2772–2781). https://doi.org/10.1016/j.pragma.2010.04.011

Stolk, A., D'Imperio, D., di Pellegrino, G., & Toni, I. (2015). Altered communicative decisions following ventromedial prefrontal lesions. *Current Biology: CB*, *25*(11), 1469–1474.

Stolk, A., Noordzij, M. L., Verhagen, L., Volman, I., Schoffelen, J.-M., Oostenveld, R., Hagoort, P., & Toni, I.

(2014). Cerebral coherence between communicators marks the emergence of meaning. In *Proceedings of the National Academy of Sciences* (Vol. 111, Issue 51, pp. 18183–18188). https://doi.org/10.1073/pnas.1414886111

Stolk, A., Noordzij, M. L., Volman, I., Verhagen, L., Overeem, S., van Elswijk, G., Bloem, B., Hagoort, P., & Toni, I. (2014). Understanding communicative actions: A repetitive TMS study. In *Cortex* (Vol. 51, pp. 25–34). https://doi.org/10.1016/j.cortex.2013.10.005

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.

Tager-Flusberg, H., & Anderson, M. (1991). The development of contingent discourse ability in autistic children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *32*(7), 1123–1134.

Toni, I., & Stolk, A. (2019). Conceptual Alignment as a Neurocognitive Mechanism for Human Communicative Interactions. In P. Hagoort (Ed.), *Human Language: From Genes and Brains to Behavior* (p. 249). The MIT Press.

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.

Warglien, M., & Gärdenfors, P. (2015). Meaning Negotiation. In *Applications of Conceptual Spaces* (pp. 79–94). https://doi.org/10.1007/978-3-319-15021-5_5