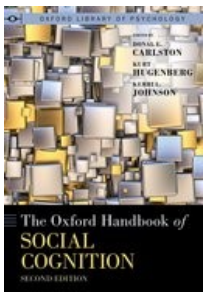

The Oxford Handbook of Social Cognition, Second Edition (2nd edn)

Donal E. Carlston (ed.) et al.

https://doi.org/10.1093/oxfordhb/9780197763414.001.0001

Published: 2024 **Online ISBN:** 9780197763445 **Print ISBN:** 9780197763414

CHAPTER

32 Language, Communication, and Social Cognition

Gregory A. Bryant, Rick Dale

https://doi.org/10.1093/oxfordhb/9780197763414.013.32 Pages 884–900

Published: 21 August 2024

Abstract

Human communication comprises many multifaceted abilities, most integrated through social cognitive systems and often manifesting in conversational interaction. This chapter draws connections between fundamental concepts in psycholinguistics, pragmatics, cultural evolution, and social cognition, addressing relationships between these domains of empirical and theoretical work. Language is a fundamental component of interpersonal communication and must interface with several related, highly interactive systems for the purpose of navigating complex social environments. The chapter describes research investigating how many nonlinguistic, social phenomena can affect language processing, revealing its deep social communicative functions. New sophisticated computational tools incorporating massive amounts of multimodal data are beginning to be harnessed in the context of interdisciplinary theoretical approaches, making big empirical questions tractable concerning the relationships between language, communication, and social cognition.

Keywords: language, communication, multimodality, conversation, alignment, pragmatics, nonverbal, cultural evolution

Subject: Cognitive Psychology, Social Psychology, Psychology

Series: Oxford Library of Psychology

Collection: Oxford Handbooks Online

Introduction

Communication is central to social life. We are born crying, we laugh and scream, and we talk in groups, often about nothing of any immediate importance and sometimes about matters of life and death. We are social animals and require social interaction much as we require food and water. Our means of communication are broad, involving all of the senses and a wealth of communication technologies. Arguably, the most powerful system of human communication is language. Through language, we can convey an uncountable variety of thoughts and feelings with high fidelity and efficient speed. Language is sometimes described as digital, in that its structure is suggestive of discrete, recombinable elements (e.g., sounds and words). These elements are subject to generative processes that afford immense communicative flexibility. Beyond this combinatorial prowess, language also derives considerable power from its deep connection with our ability to understand communication in social contexts.

The origins of our complex linguistic abilities are not well understood, and to make matters more complicated, language operates alongside other communication systems such as facial and vocal signaling and within the context of a dynamic, interactive world. In particular, language *is* social. It is inextricably intertwined with conceptual and social knowledge. Many languages even encode socially relevant factors
p. 885 such as politeness directly into their grammars ↳ (e.g., Helmbrecht, 2003). The evolutionary connection between social cognition and communication is also evident in the apparent neural homologs of orofacial expressions in monkeys and human speech (Shepherd & Freiwald, 2018). The origins of language and speech are rooted in the social communicative behaviors of our primate ancestors, underscoring the profound link between social cognition and communication (Fitch et al., 2010).

In this chapter we aim to describe some approaches to language and social cognition that we argue offer a framework for bridging these domains. It is impossible for one review to cover all prior work relevant to these areas. However, the selective review here maps both the empirical and the theoretical landscape so that readers, we hope, come away with a sense of emerging directions of important research in language, communication, and social cognition.

Language and Cognition

Scholars have debated for decades on how to define language. Attempts range from listing multiple proposed design features (Hockett, 1960) to reducing language to a single computational mechanism (Chomsky, 1995). More recent conceptions regard language as a culturally evolved, construction-based skill rooted in general cognitive abilities (e.g., Christiansen & Chater, 2008) or as a collection of adaptations solving a suite of language-specific problems (e.g., Pinker & Jackendoff, 2005). Debate often centers on the evolutionary origins and functions of language and its relationship to communication. To what extent have linguistically relevant cognitive traits, such as increased memory and attention, changed or evolved to serve in a communicative capacity? Perhaps cognitive mechanisms like memory and attention were integrated for language only recently, and each mechanism has its own unique, nonlinguistic evolutionary history (e.g., Carruthers, 2002). Another intriguing possibility is that language takes the form that it does because it has adapted to the cognitive constraints of the human brain, rather than the brain adapting for language (e.g., Christiansen & Chater, 2016; Kirby et al., 2014).

Any approach to the nature of language has entailments for characterizing its relationship to other aspects of our psychology, including social cognition. We consider language to be, in the broadest sense, a culturally evolved, combinatorial cognitive system by which people communicate in complex multimodal contexts. Language interfaces dynamically with most other aspects of cognition, and it relies heavily on the ability of receivers to draw rich inferences about the meaning and intention of others (Gibbs, 1999; Scott-Phillips,

2015). Our aim here is to illuminate some interesting elements of the interface between communicative action mediated by language and social cognition more generally. First, we address an issue of what is called the modularity of language. This situates our discussion here in broad theoretical terms, because the classic question of whether language behaves like a modular system impacts our understanding of the connections between language and social cognition.

On Modularity of Language and Social Cognition

The topic of modularity is both historically and theoretically important for understanding the relationship between language and social cognition. The concept of modularity found its strongest expression in the work of Fodor (1983), who described many cognitive processes, including language, as being specialized subsystems. According to this framework, these subsystems, or *modules*, have specific properties such as being encapsulated (i.e., highly independent, self-contained processes), innate, automatic, fast, and so on (see Pylyshyn, 1999, for discussion). Fodor's proposal has had a great influence on many programs of cognitive research, especially in the perceptual sciences, and has inspired considerable theoretical debate (Colombo, 2013).

p. 886 The concept of modularity has implications for social cognition. If language is an encapsulated subsystem separate from other mental processes, then it follows that social cognitive processes would not play a role in language. But if social cognition (such as social perceptions or judgments) influences language in some way (e.g., how it is produced or understood), does it mean linguistic subsystems are *not* encapsulated? In fact, much research suggests that language is not encapsulated and that it can be influenced by other processes, such as perceptual or social psychological mechanisms. For example, some research has shown that speech processing in the brain is modulated by the identity of a speaker. If you hear a stereotypically male voice mention a stereotypically female topic, the brain shows a rapid response indicating that language and social inferences might be happening simultaneously to process the inconsistent information (Van Berkum et al., 2008). Moreover, many behavioral and neurocognitive studies have shown that perceptual systems can be affected by other kinds of cognition, including language (for a brief review, see Vinson et al., 2016).

Many cognitive and perceptual systems have a close relationship with language production and understanding. Research demonstrating this abounds, using brain imaging (R. Tomasello et al., 2019), measures of behavioral dynamics (Falandays et al., 2020; Freeman et al., 2011), psychophysical experiments (Lupyan et al., 2020; Lupyan & Ward, 2013), and more. Research in the burgeoning area of "social vision" has also demonstrated how social processes, such as person or situation judgments, may directly impact processes of vision and other sensory or perceptual systems (Dunning & Balcetis, 2013; Johnson & Shiffrar, 2013). Taken together, these findings support the idea that language, communication, and social cognition constitute mental processes that are more integrated than the classical concept of modularity implies. Issues are still debated, especially in perception (e.g., Firestone & Scholl, 2016) and in various domains of the language sciences where Fodor's influence still echoes (e.g., Chomsky, 2017; Friederici, 2011; Kingston et al., 2016; Norris et al., 2018).

Though many researchers have focused on specific Fodorian conceptions of modularity, others have developed more nuanced approaches. For example, taking an evolutionary approach, Barrett and Kurzban (2006) proposed that theoretical emphasis should be on functional specialization in cognitive systems and not on strict conditions of modularity per se. Rather than follow a prescriptive list of properties, they suggest we should look for design features of cognitive systems that are shaped by natural selection to solve particular tasks. In this case, language does not get defined as a "module" a priori under strict criteria, but instead researchers explore aspects of language that might represent adaptations in a larger network of cognitive, social, and behavioral characteristics. The list of potential adaptations is long: theory of mind, vocal production and perception, flexible symbolic gesturing, and so on (cf. Hurford, 2003).

Relatedly, other researchers have turned to cognitive neuroscience. Whatever language or social cognition may be in our brains, the best way to find out is to investigate the brain itself. A network approach to brain organization provides a framework for evaluating interactions we see in language and social cognition (e.g., Anderson, 2014; Bullmore & Sporns, 2012; Fedorenko & Thompson-Schill, 2014). By this view, the brain's networks are dynamic and respond to different tasks and situations. Language and social cognition are not taken to be strictly separate subsystems, but rather outcomes of these brain networks. This research suggests that language may indeed be implemented in functionally specific regions in the brain, but these regions are part of a highly interactive network that does not coincide with traditional conceptions of modularity.

These observations suggest modularity is a continuous, relative property, not all or none. Clearly,
p. 887 perception is specialized to the extent that different modalities require certain kinds ↳ of receptors to extract information—ears cannot process light, and eyes cannot process sound. But beyond this specialized hardware, at higher neural levels, discovering details of relative cognitive specialization (or nonspecialization) characterizes recent discussion (cf. Fedorenko & Thompson-Schill, 2014). Researchers are continuing to investigate how such a collection of massively integrated information processors do their work. In the case of language and social cognition, it helps to look at the behavioral context in which these systems have been shaped to operate in: conversational interaction.

Language and Conversation

The issues summarized in the prior section are about the mental structure of language and so relate to conversation, the primordial ecology of language use. Conversational interaction is an inherently social activity. We effortlessly engage in talk, and most of the time we understand one another reasonably well. But the processes underlying that interaction, including the production, perception, and understanding of language and speech, are vastly complex. As conversationalists, we must integrate multiple sources of information, including visual, auditory, and even olfactory input during engagement. Simultaneously, we prepare upcoming sentences, produce feedback signals, and regulate our interpersonal behavioral timing in fine-structured ways. On top of the specific interactive problems we solve, we are engaged in social action. H. H. Clark (1996) described language interaction as a type of joint activity. People in conversation generally share communicative goals and establish common ground, meaning that we develop a sense of how our knowledge, desires, and immediate goals align with those of our fellow interlocutors (these ideas also tie into what is called *Gricean pragmatics*, discussed further in the section Social Pragmatics. Pickering and Garrod (2004) suggested that this interactive alignment unfolds dynamically and that various linguistic and nonlinguistic behaviors update and modulate the alignment stability. Discovering the underlying processes that steer us in this complexity has been central to theoretical debate (Barr, 2014; Brennan et al., 2010; Brown-Schmidt, 2009; Dale et al., 2013; Pickering & Garrod, 2013; Shintel & Keysar, 2009).

The systematic study of conversation has been subject to a variety of methodological traditions. But researchers who agree that language and its social/situational milieu cannot be sharply demarcated often differ on the best means by which to study this integration. Many important influences on language processing have been shown in the laboratory. These studies tend not to focus on natural language use, but rather examine "textoids" that characterize brief segments of text (Graesser et al., 2007), which Garrod and Pickering (2004) referred to as the study of "monologue." A number of highly replicable and stable findings about how the mind processes language have been discovered through measuring how fast participants perceive and comprehend short segments of text, even words. A simple but classic example is the effect of frequency, related to the social cognitive domain of concept accessibility. Words that are more frequent tend to be recognized more quickly, and sentences that use common phrasings tend to be read more smoothly (Balota et al., 2007).

To understand the relationships between communication, cognition, and language in common use, research methods must be developed that investigate conversation holistically, embracing its multimodal and dynamic complexity. Some researchers have encouraged highly qualitative and focused transcript and audiovisual analysis, such as in the tradition of conversation analysis (C. Goodwin & Heritage, 1990; Heritage & Clayman, 2011). For example, by carefully describing and coding samples of talk, conversation analysts have found that human interaction is a highly local phenomenon—interlocutors often respond to each other promptly in a kind of chain of adjacent contributions. The simplest example of this is a ↳ greeting, but other adjacencies include a host of familiar conversational pairs, such as query–answer, inform–confirm, and so on.

One concern with conversation analysis is that it is still developing a solid quantitative basis. And so other scholars have developed large-scale transcript analysis in a subdiscipline of linguistics called *corpus linguistics*. In this work, researchers obtain quantitative measures of factors associated with aspects of discourse. For example, the work of Biber and others has analyzed many thousands of samples of conversational transcripts to determine how often disfluencies occur, which phrases are most common, which words tend to be used repetitively, and so on (e.g., Quaglio & Biber, 2006). The result is an aggregate quantitative description of the structure of interaction.

Procuring this statistical description in corpus linguistics is powerful, but it lacks experimental control, limiting causal inferences about what drives language and communication. Other researchers have expanded the tradition of psychological experimentation to create unscripted yet controlled interactive designs, to see how social factors may influence language and communication (e.g., Brennan et al., 2010). For instance, in a classic experiment sometimes called the *tangram task*, participants work together to refer to unfamiliar shapes. By observing how participants do this, experimenters can observe (and manipulate) what communicative strategies they use. One key finding in this work is the rapid rise of efficiency: Participants coordinate with each other to devise something akin to their own local language system just for the purpose of referring to these unfamiliar objects. Some researchers have expanded measurement and analysis in these natural contexts. By measuring millisecond-level dynamics of various behaviors in these interactions, such as eye tracking or body motion, researchers can obtain the kind of temporal precision characteristic of traditional laboratory designs mentioned above (e.g., Cornejo et al., 2017; Fusaroli et al., 2016; Paxton & Dale, 2013). This will be described further in the section Multimodality, Embodiment, and Dynamics.

Language and Human Sociality

We will now examine the overarching framework that underlies language and social behavior—why we engage with one another using language at all—and that is cumulative culture and the evolution of cooperation.

Cumulative Culture, Language, and Cooperation

Human communication necessarily manifests itself in a cultural context, and cultural evolutionary processes clearly play a role in how languages evolve and how they are used. We borrow a definition of culture from Dean et al. (2014), which is "group-typical behavior patterns shared by members of a community that rely on socially learned and transmitted information" (p. 285). Many animals exhibit behaviors that qualify as culture by this description, but most scholars agree that humans are unique in our propensity for retaining transmitted information and building on it. M. Tomasello (1994) described this cumulative process as a ratcheting effect whereby cultural products can become increasingly complex to the point that no single individual could have invented such a product de novo. Cumulative culture is a result of cultural evolution—that is, processes of variation, inheritance, and selection that we see in biological evolution have analogues in the processes that preserve accumulated changes to cultural information over time. Examples abound in technology, fashion, art, and just about every aspect of human life—we can see the slow improvements and varying modifications on human artifacts that unfold across generations. A famous example is found in changes in stone tools in archaeological studies. Distinct sequences of accumulating effects can be tracked over long stretches of time, including changes in style and effectiveness (Ambrose, 2001). Writing ↳ systems provide another classic example, where protowriting appears to be
p. 889 specific to city-state financial administration (i.e., numerical notation) and slowly evolved in stages, and in different locations, to fulfill highly diversified functions requiring the representation of linguistic information (Mesoudi et al., 2013). More fundamentally, the transformation of any cultural product over time is the result of similar processes, yet the origins of most cultural phenomena are largely unknown to the majority of people who engage in them at any point in time.

There exists much debate regarding the usefulness of a Darwinian framework for understanding cultural evolution, as well as what components of our evolved psychology afford cumulative culture. These debates are beyond the scope of this chapter, but certain safe assumptions can be made: (a) human culture is rooted in social cognition, (b) language plays a key role in facilitating cultural transmission, and (c) culture evolves at some level in the service of cooperative interactions. Cumulative culture, language, and cooperation constitute a good part of what makes us human. But what is cooperation?

Cooperation can be characterized as an interaction resulting systematically in a positive outcome for all individuals involved. The positive outcome is a result of behavioral adaptations that reliably generate it, and not occurring by chance. This definition avoids including situations where individuals in a group incidentally benefit (a byproduct) from an interaction, despite underlying behavioral causes unrelated to a positive *mutual* outcome (e.g., two nonhuman animals simultaneously pursuing the same prey animal). The approach also requires that individuals behave cooperatively in every instance of it, even if that behavior imposes an incidental immediate cost or does not result in a positive outcome in some cases. Put simply, cooperation is the situation where two or more individuals act, by design, in a way that generates mutual benefits *on average*. For human sociality we are interested in the variety of cooperation that evolves between unrelated individuals. From the perspective of comparative biology, we have a framework from which behaviors can be effectively compared. By this view, there are no widely agreed-on cases of cumulative culture in a nonhuman species. Cooperation, as defined here, is rare, but does occur under the right conditions in a variety of organisms.

In humans, spoken and written language is the primary vehicle by which cumulative culture and cooperation occur. The structure of language affords accurate transmission of information with very little cost. We share information, groom one another, and declare intentions, all in the service of cooperative interaction. The only reason we can manage these kinds of interactions is because we process rich social information informing our expectations about others' behavior. We have to know what others are going to do, and why. Complex social life requires a deep appreciation for the perspective of others. This requires perspective taking, central to social cognition. We must imagine what people are feeling, how those feelings

might affect their behavior, and what the world looks like from their perspective. Not only can we imagine what others perceive, but we can also imagine what they are thinking and simulate various mental states. Researchers have described this ability as metarepresentation and theory of mind. The ability to attribute distinct thoughts, beliefs, and intentions to others emerges early in development (~12 months) and unfolds in tandem with linguistic and social skills. For example, Surian et al. (2007) tracked the visual attention of infants while they viewed animated videos of agents and found that infants' attention was consistent with an awareness of the internal knowledge and belief states of the characters in the animation.

From a communication perspective, multiple systems come online in massively interacting, dynamic ways, resulting in coordinated facial and vocal signals, gestures, and other body movements. These systems are bootstrapped by cognitive processes, perhaps especially related to social understanding. The foregoing discussion on culture and cooperation illustrates ↳ this—further illustration of the relationships between social cognition and language is found in research on how children acquire language.

p. 890

Language Learning, Grounding, and Social Cognition

Perhaps the most convincing linguistic universals pertain to how language is acquired. All children learn their first languages in the midst of other humans, interacting or observing interaction in rich social and physical contexts (E. V. Clark, 2009; Oller, 2000). No children learn their first language(s) primarily from televisions, computer programs, or operant chambers. While early learners tend to acquire language according to a particular sequence, this can vary both across children learning the same language and across languages. Consider word learning, for example. Children tend to acquire words early that pertain to concrete objects or actions (Gilhooly & Logie, 1980; cf. Łuniewska et al., 2016) and learning is linked to situational factors, such as the specific word context (e.g., Roy et al., 2015). The social interactive factors at play are currently under debate, but there is abundant evidence that children and their caregivers are in a dynamic interplay that supports this learning. In both laboratory experiments and natural recordings, researchers have found that the dynamic structure of interaction, such as turn taking and verbal or nonverbal responsiveness, facilitates word learning (e.g., Warlaumont et al., 2014; Yu & Smith, 2012).

These structured aspects of language acquisition accompany the social elements of communication that are also rapidly learned. In fact, the capacity to interpret the goals of communication partners emerges quite early in development (before 6 months: Csibra & Gergely, 2009) and may constitute a domain-specific cognitive specialization (Cohen & German, 2010). Correspondingly, as noted above, the capacity for children to engage in shared, goal-directed visual attention may be unique—or at least most pronounced—in our species (Warneken et al., 2006). This attention may be important for language learning and social cognition *together*, suggesting that these two major aspects of human cognition are developing in parallel (for debate and discussion, see Akhtar & Gernsbacher, 2007; Mundy & Newell, 2007).

For example, consider very early language-related learning, such as the control of speech and vocal expression. Observational studies of very early vocalizing by children demonstrate significant flexibility in how their speech corresponds with particular social goals, such as cries of distress or expressions of interest (Oller et al., 2013). The argument is that some degree of decoupling between the speech capacity and specific functions and contexts permits flexible learning so that complex language may emerge from exploration (Ritwika et al., 2020). Thus, humans might be equipped with both species-specific vocalizations prepared for particular emotional and social signaling (such as laughter and crying) and vocal behavior that gets tuned to new and distinct social functions involving relatively more learning. This learning may require social and environmental input to shape culturally relevant language behavior.

Historically, there has been much debate regarding the nature of language input to children, and the issue links importantly to the topic of modularity discussed above. Some researchers have considered language,

or some aspects of language, to be a kind of innate module. By this account, language is prescribed genetically and simply matures amid language input much like the way other organs of the body mature in response to various external effects. For decades, following Chomsky's early and influential proposals for innateness many researchers have identified social and environmental factors that can contribute to language development, such as frequency, word-learning context (Roy et al., 2015), and parental feedback (Tamis-LeMonda et al., 2014). Some of this work has made use of corpus linguistic methods, analyzing how children start to use language and whether transcripts of interactions can predict these tendencies at various levels (from ↳ words to syntax, etc.; for review, see Ambridge & Lieven, 2011; Christiansen & Chater, 2016; M. Tomasello, 2009). This also pertains to social cognition, as noted above, because the learning mechanisms for language may involve following the attention of caregivers in context, coming to understand the goals of communication, receiving rich structure from the social environment in various ways, and so on. In general, it seems likely that both innate and environmental factors contribute to language learning, but further discussion is outside the scope of this chapter.

p. 891

Communication and Language Use in Context

We now turn to how people use language in social contexts. Rather than viewing language strictly as a code that contains information transferred between people, scholars have developed approaches that instead recognize language as one important source of evidence for intended meanings situated in rich social environments.

Social Pragmatics

During the middle of the 20th century, philosophers of language made strides toward understanding the utility of language as a social tool. Rather than focus on the details of the formal properties of linguistic structure, scholars began to describe how people use language to achieve social goals. Austin (1962) famously described how people do things with words. Language production, from this view, can be construed as speech acts that have real effects in the world, whether by the mere act of saying something (e.g., "I now pronounce you husband and wife") or by causing others in the world to recognize an intention, belief, or other action. Austin proposed that language use exists on at least three levels: locution, illocution, and perlocution. Locutionary acts correspond to the surface features of utterances—what people literally say. Illocution refers to speakers' underlying meaning, or intent; and perlocutionary effects describe behavioral (or cognitive) outcomes in listeners. This approach to language put a much greater emphasis on the importance of receiver adaptations in which social beings interpret people's surface communicative behaviors through inference. As described later in this section, this idea evolved into the theory of ostensive communication, a backbone of social communication (Scott-Phillips, 2015).

Grice (1975) described how speech acts resulted in meanings he called particularized implicatures—unstated propositions associated with the purposeful flouting of conversational maxims. Maxims can be thought of as implicit rules that conversationalists follow that facilitate mutual understanding—guidelines based on an assumption that interlocutors are cooperative communicators. This is a different notion of cooperation than described earlier, as we explain in the section Pragmatic Strategies and Processes. According to Grice, speakers aim to be appropriately informative (quantity), truthful (quality), and relevant (relation) and to do so in a way that is maximally clear and unambiguous (manner). By violating (i.e., flouting) one or more maxims, speakers intentionally communicate an alternative meaning from what the literal words might otherwise suggest. For example, in response to the question "How was the movie?" a respondent could utter "The popcorn was good." In flouting the maxim of relation by not referring to the movie specifically, the respondent communicates a negative opinion of the movie. But in other cases of

implicature, flouting is not needed for listeners to derive implied meaning. Generalized implicatures do not require specific contexts to be understood, and are inferred because listeners assume speakers *are* following the maxims. For instance, the statement "John walked into a house a saw a tortoise" implies that the house was not John's. If it were his house, this information would be expected in the sentence meaning (i.e., maxims of relation and manner).

p. 892 Austin (1968) and Grice (1975) opened the door for language researchers to explore the ways that people use language strategically to communicate a variety of meanings. People clearly say things that are quite different from the exact meanings they wish to convey, but until early pragmatic theorists proposed such multilayered elements in linguistic production, there was no account of how language actually worked beyond the generation of surface linguistic units. One crucial empirical problem for Grice's (1975) account is that it made rather specific predictions regarding the time course of cognitive processing in language understanding that do not bear out in experimental studies. In general, it should take longer for listeners to process meanings that require a recognition of maxim flouting than for literal meaning not requiring that step, all else being equal. Studies examining the online processing of indirect speech, for example, have shown that in many cases, listeners acquire indirect meanings as fast as or even faster than they do for the literal counterparts to these utterances. In psycholinguistic tasks in which participants are timed in how quickly they comprehend literal versus nonliteral messages, researchers often find that there is no added cost to using nonliteral phrasing—it can be processed just as quickly as literal phrasing (for review, see Gibbs, 1999). Nevertheless, Grice (1975) offered an account that addresses an empirical challenge to language researchers. How do people arrive at their interpretations of meaning when processing language that has no literal connection to what speakers mean? At its core, this is a problem of social cognition, even if Grice and other early language philosophers did not conceptualize it that way.

Sperber and Wilson (1986/1995) reduced Grice's maxims to an overall principle of relevance as part of a larger theory of ostensive communication. On this account, communicative acts are composed of two layers: a communicative intention and an informative intention. A communicative intention informs a receiver of an intention to communicate something relevant, and the informative intention is conveyed through linguistic and other evidence, with the goal of changing the information state in that receiver. Infant-directed speech offers a nice example. In many cultures, when speaking to infants, speakers often produce modified speech, including features such as raised pitch, noticeably raised or lowered loudness, lilting rhythms and intonation patterns, and a variety of lexical and other linguistic characteristics that are attractive to infant listeners (Cox et al., 2022; Fernald, 1989; Hilton et al., 2022). The unique speech patterns provide a signal of communicative intentions, meaning that target listeners realize that the message is intended for them. In the case of infant-directed speech, the content is often irrelevant and the communicative intentions are the entirety of the message. But speech directed at particular targets can have any number of special features that make it especially pertinent and noticeable to them—that is, makes it relevant to their cognitive environment (Sperber & Wilson, 1986/1995).

By producing an utterance with special properties, speakers can make language easier for listeners to process. According to this theory, interlocutors seek optimal relevance. Doing so maximizes the cognitive benefits of communicative acts and minimizes the cognitive effort to extract relevant meaning. For example, interaction partners tend to converge quickly on the words or phrases used to communicate about complex or novel concepts. By maintaining a consistent terminology, even if it's just in that moment of conversation, interlocutors perceive the continued relevance of their dialogue and expand the conceptual range of their discussion while facilitating each other's mental processing (e.g., H. H. Clark & Schaefer, 1987). Moreover, speakers will produce hesitations, approximators, so-called disfluencies (e.g., um and uh), and various prosodic signals to assist listeners in even the most basic exchanges. One study of how people answered questions across minimally varying pragmatic situations found that slight contextual changes can have predictable effects on responses. Gibbs and Bryant (2008) asked people in the street to give the time,

p. 893 with several variations in the question, ↳ ranging from simple requests ("Excuse me, do you have the time?") to more detailed ones ("Excuse me, I have an appointment at [some 30 min interval]. Do you have the time?"), and found that people's answers contained many speech features likely produced in direct response to implied information requests. For example, people predictably gave exact time versus rounded responses and approximators such as "about" or "around" depending on an implied need for precise time estimates. Responders additionally used signals like "umm" and "uhh" and often took extra time, indicating they were processing the implied request. This work revealed, in multiple experiments, that people make communicative efforts to accommodate others' implied information needs and that social cognition drives language production, measurable in the dynamics and content of spontaneous speech in response to simple requests by strangers.

Pragmatic Strategies and Processes

Language use is grounded in the dynamics of everyday social interaction. As mentioned above, we can conceptualize the ways that people deploy linguistic tools during social communication as strategic. This is not to say that people are aware most of the time of their strategizing, but rather that social interactions involving language can be conceptualized in game-theoretic terms. By exploiting the distinction between what is said and what is meant, language users can imply rich meanings for various gains while minimizing the potential costs of those unstated meanings. The concept of implicature, described earlier, is often illustrated with common examples like indirect requests. A speaker may imply that a listener (seated next to a window) should do something when the speaker says, "It's stuffy in here."

Historically, these aspects of interaction were characterized as a kind of cooperative behavior. But even if language users are following pragmatic rules for good communication, they could be engaged in an exploitative interaction. An alternative meaning of cooperation comes from evolutionary biology. In this sense, we could see even seemingly exploitative communication as having a positive outcome for all individuals in a group. These positive outcomes result from behaviors adapted to reliably generate that outcome repeatedly. For example, inducing a conversation partner to carry out some chore after an indirect request benefits the speaker by having the request fulfilled, but it may also benefit the listener by accommodating their conversation partner and signaling a willingness to help (even if it is done unconsciously or under some social coercion).

Consider how introducing biological cooperation dynamics into language use alters our understanding of how language interfaces with social cognition. Pinker et al. (2008) described the case of the rational briber. Imagine a situation where Harry is being pulled over by a police officer for speeding. During the interaction, while the officer decides whether to issue a ticket, Harry can ask if it is possible to take care of the ticket then and there, while making a $50 bill visible in his wallet, which he is handling to provide his identification. The officer notices this and can (a) accept the implied bribe, (b) accuse Harry of bribery, or (c) ignore the move altogether and proceed with giving him the ticket. The problem with an accusation is that there is no direct evidence, such as an explicit offer of $50 in exchange for the ticket being thrown out. This really only leaves the officer with two choices: be an honest cop who issues a ticket and ignores the bribe or be a dishonest cop who takes the money and sends Harry on his way. Harry has three choices in this scenario: (a) don't bribe and almost certainly get a ticket that could be $(*x*), (b) directly offer $(<*x*) as payment to be let go, or (c) indirectly make an offer (i.e., $<*x*) that has plausible deniability. Strategically, there is an obvious best choice for Harry: use indirect speech.

p. 894 The scenario described above can be laid out in a payoff matrix in which potential costs and benefits of each outcome can be compared. When using indirect speech, people can implicate meanings that are officially uncertain, substantially reducing risk in a variety of scenarios as a result of plausible deniability. But it still allows for certain users to access meaning that could be mutually beneficial. In our example, dishonest cops

like bribes, but Harry does not know whether he is dealing with a dishonest cop. If he is not, the cop can at worst only suspect a bribe took place, but cannot do anything about it legally. This logic extends to many situations, including veiled threats, sexual come-ons, dog whistles, ironic criticism, and so on. Classic examples include threats using the form of "it would be a shame if …" and then stating the threatened action or sexual come-ons where an individual suggests a friendly interaction with an implied intention for intimacy (e.g., after a dinner date, a person asks, "Would you like to come upstairs for a nightcap?"). In both scenarios, there is plausible deniability of the implied meaning. In the case of the veiled threat, the sender is not literally stating one and so cannot be accused of such with any legal validity. In the sexual come-on example, the sender is *possibly* being sincere in their platonic intentions, thus preserving the relationship in case the advance is unwelcome. In all cases, senders exploit ambiguity to avoid potential costs, but maximize possible rewards.

The costs in these situations vary according to the different social scenarios where these dynamics can play out and the social roles and power dynamics of those involved—even in a mundane situation like trying to bribe a maître d' at a restaurant to get a quick table. Direct requests, or demands, can usurp power roles and thus are subject to rejection by those who, at some level, do not wish to surrender their power. Indirect requests allow others to save face but still grant requests. There are many reasons that we cannot quite say what we mean because it could be costly to our reputation, damage a relationship, or even break the law. Thus, we imply it strategically, and if done correctly, the act can afford all parties involved some kind of benefit. The domain of politeness has offered empirical evidence for the importance of these strategies, such as how the social identity of a conversation partner may complexly alter communication practices (Idemaru et al., 2019). It is cooperative in the biological sense, as we defined earlier.

But what cognitive processes underlie such communicative dynamics? Strategic social communication often must involve language, because there are few alternatives to indicate a meaning that implies an unstated intention. Language users afford specific interpretations of their utterances by using devices that are relevant for hearers, but we must have some kind of cognitive ability that allows us to process the complexities of multiple meanings. Moreover, this ability must interface with systems that help us recognize the mental states of other actors in a given scenario. We must have, at some level, a model of other minds that guides our understanding of intentions. Theorists have described this phenomenon in a variety of ways, including metarepresentation (Sperber, 2000) and theory of mind (Perner, 1991). That is, we can represent the beliefs, desires, and intentions of other people as distinct from our own mental states. Language users exploit this by generating surface linguistic features that correspond to one intention, or mental state, and appeal to an understanding of an alternative, implicit state of affairs. For example, a large literature documents the important relationship between metarepresentation and verbal irony use, including clinical neuropsychological research revealing brain deficits associated with language understanding impairments (for a review, see Bryant, 2012). In the case of irony, speakers and listeners must understand each other by recognizing that a surface description (e.g., "Another lousy day!") deliberately and humorously contrasts with a current state of affairs (e.g., sitting outside in perfect weather).

One recent approach is to examine language understanding as a probabilistic process of decision under
p. 895 uncertainty. Many sources of information must be integrated by language users ↳ in an effort to understand speakers' intentions, and over time certain expectations build based on experience and social knowledge. Inspired by the Gricean notion of language use as rational action, rational speech act theory formalizes comprehension, resulting in models that predict people's interpretations of all language use, including indirect language such as verbal irony and metaphor (Goodman & Frank, 2016). Speakers are assumed by listeners to strive for utility maximization and be informative, but these functions can be modified to adjust reasoning strategies in different contexts and follow a game-theoretic logic. The rational speech act model represents an important advance in the formalization of pragmatic processes.

Multimodality, Embodiment, and Dynamics

All of the aforementioned pragmatic strategies are embedded in a very complex array of communicative behaviors. The multimodal nature of ordinary human communication has inspired a number of recent programs of research. Multimodality, as noted above, is the idea that any language performance involves the coordination of many different behaviors and cognitive processes. In face-to-face communication, for example, interlocutors must simultaneously produce and perceive gestures (forms and dynamics), prosodic modulation (e.g., pitch, loudness, and speech rate), social gaze and other aspects of visual attention, and so on. The term multimodality is also sometimes used to refer to the many types of information levels, such as abstract linguistic features of dialogue, word choice, phrasal preference, conversational topic, and more. Speakers and listeners must also contend with the probabilistic nature of these signals, including when they conflict, such as when gesture and speech are misaligned (Kelly et al., 2015).

Much interest in embodiment stems from experimental studies showing how language processing is impacted by perceptual or motoric information, harkening back to our discussion of cognitive penetrability. In one line of work, researchers have shown that participants may "simulate" aspects of a text in their minds, and these simulations contain spatial and perceptual ingredients. Spivey and Geng (2001) found that while participants imagined objects that were not in a visual display, they moved their eyes in a manner consistent with spatial aspects of the provided context. For example, a story centered around a high rise had participants gradually moving their eyes upward as if they were simulating being present in the story. This is just one of many such demonstrations, and they have been the subject of significant debate. Some researchers contend that these embodied characteristics only capture a small part of our language and communication capacities (see discussion in Zwaan, 2014).

In the past, many research strategies have involved tasks that are unnatural, such as laboratory experiments. These experiments often have participants interact in some structured but unscripted task, such as an interactive naming game in which participants refer to or remember novel objects (Dale et al., 2011). Such experiments offer excellent control of variables, but ultimately lack substantial external validity. Other researchers have sought embodiment and multimodality in natural contexts, including in qualitative analysis of language samples from recordings in natural situations, such as at home (VanDam et al., 2016). These samples are definitively natural and ecologically valid. This robust multimodality of natural human language and communication has inspired those in conversation analysis to focus their attention both on the linguistic (verbal) form and on how nonverbal accompaniments help to structure the creation of meaning. This work is related theoretically to research on social contagion by Chartrand and others (Chartrand & Lakin, 2013). For example, M. H. Goodwin (2017) and C. Goodwin (2017) argued that interactions among co-present individuals involve dynamic embodied engagements, something M. H. Goodwin (2017) has called "haptic sociality." Research in this domain has also looked at speech and gesture timing together. This work ↳ typically uses qualitative, fine-grained inspection of carefully coded conversational transcriptions to study coordination (for a review, see Alviar et al., 2023). In more quantitative work, recent generation of multimodal corpora reveal rich patterns of coordination both within and across communicating individuals (Alviar et al., 2019; Fusaroli et al., 2016; Louwerse et al., 2012; Pouw et al., 2020). In this research domain, a major effort is first invested in devising methods to automatically or semiautomatically produce multimodal transcripts that contain rich information about the timing of modalities.

p. 896

This approach to language may lead to future integration with those reviewed in the prior sections. Theories that involve social pragmatics—which may be termed "high-level" theories because they invoke rich cognitive processes and strategy—involve elaborate processing implications for human minds in interaction. Indeed, sometimes the term *contract* is used to describe underlying cognitive commitments implied in these theories (Tollefsen & Dale, 2012). By contrast, research drawing on multimodal dynamics has anchored more to "lower level" theories, based on rapid dynamic processes. In this account, two people

communicating have direct causal impact on each other via perception and action and need not always invoke high-level cognitive strategy (cf. Shockley et al., 2009). The various traditions and empirical approaches described across this section reflect important ongoing research to determine the aspects of cognition (including social cognition) that underlie our capacity for such rich social pragmatics *and* a veritable orchestra of multimodal processes while interacting (e.g., Barr, 2014; Brennan et al., 2010; Dale et al., 2013; Fusaroli et al., 2014; Shintel & Keysar, 2009). Bridging these high-level processes, such as communicative strategy, to low-level processes, such as perception and action, is an important frontier domain of this discipline.

Conclusion

In this chapter, we have summarized a variety of approaches to and connections between language, communication, and social cognition. Some important history was shared, such as classic approaches rooted in the cognitive science of language, from the modularity and innateness of language to the rich social environment in which language is learned. We covered other historically important perspectives, such as using Gricean principles to formulate a new basis for understanding intent and cooperation during interaction. In the previous section, we illustrated an important next step across these research traditions: to move beyond simplistic theorizing and find important connections between concepts of communication rooted in different traditions. Understanding *how* we manifest our rich communicative signals ("from blinks to winks," Riley & Turvey, 2001) and *why* we do so for various social strategies is very much a research area waiting for synthesis. It is generally recognized that embodiment and other multimodal dynamics are critical for language and communication (H. H. Clark, 2021; Goodwin, 2017), and this domain is ripe for new theoretical generalizations to articulate these relationships, if possible in formal terms.

Language is a communicative tool that is shaped by evolved cognitive mechanisms and the forces of cumulative culture. It is a deeply social phenomenon that interacts with many other cognitive systems. Moreover, language evolved to occur in multimodal, highly social interpersonal contexts that include a vast number of other communicative systems, some evolutionarily much older, such as nonlinguistic vocal behavior, body gestures, and facial expressions. Scholars have grappled for decades with how to carve these cognitive and communicative abilities at their joints, but it is becoming increasingly clear that most aspects cannot be properly understood empirically independent from one another. For example, the cultural evolutionary processes that shape language are, in many ways, themselves largely fueled by language, with
p. 897 multiple systems in a complex feedback loop. Our social reasoning and language are deeply ↳ intertwined, with some scholars claiming that reasoning evolved, at least in part, to facilitate persuasive language-based argumentation (Mercier & Sperber, 2011).

Research to date has brought us to this point in our understanding, and there is now a robust recognition of the vastness of the empirical problems at hand. Consequently, empirical and theoretical approaches have matured, as have our technological tools for conducting research. Computational social science techniques involve amassing huge amounts of data—researchers face the difficult task of developing proper algorithms and statistical approaches to analyze and understand that data. We are now able to document everyday interactions of people across widely diverse cultures and contexts in high-resolution, real-time, multidimensional formats (video, audio, physiological, etc.). The sophistication of our methodological approaches needs to match that of our amazing new tools. Interdisciplinary research is increasing rapidly, integrating theoretical and empirical tools across all sciences. Here, we have presented a view toward these complex issues from the perspectives of cognitive science and evolution. These complementary approaches afford theoretical integration for describing and explaining multiple levels of analysis in linguistic and social interaction. But a complete understanding of language, communication, and social cognition will likely require analytical approaches using computational tools not yet developed or incorporated into the

social sciences. Nevertheless, the future is bright as we move toward an understanding of our most basic human abilities through the use of those same powerful skills that we embody so fundamentally.

References

Akhtar, N., & Gernsbacher, M. A. (2007). Joint attention and vocabulary development: A critical look. *Language and Linguistics Compass*, *1*(3), 195–207.
Google Scholar WorldCat

Alviar, C., Dale, R., & Galati, A. (2019). Complex communication dynamics: Exploring the structure of an academic talk. *Cognitive Science*, *43*(3), e12718.
Google Scholar WorldCat

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

Ambridge, B., & Lieven, E. V. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Ambrose, S. H. (2001). Paleolithic technology and human evolution. *Science*, *291*(5509), 1748–1753.
Google Scholar WorldCat

Anderson, M. L. (2014). *After phrenology* (Vol. 547). MIT Press.
Google Scholar Google Preview WorldCat COPAC

Austin, J. L. (1962). *How to do things with words*. Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445–459.
Google Scholar WorldCat

Barr, D. J. (2014). Perspective taking and its impostors in language use: Four patterns. In T. Holtgraves (Ed.), *The Oxford handbook of language and social psychology* (pp. 98–110). Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review,* 113(3), 628.
Google Scholar WorldCat

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 53, pp. 301–344). Academic Press.
Google Scholar Google Preview WorldCat COPAC

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, *16*(5), 893–900.
Google Scholar WorldCat

Bryant, G. A. (2012). Is verbal irony special? *Language and Linguistics Compass*, *6*, 673–685.
Google Scholar WorldCat

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–349.
Google Scholar WorldCat

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, *25*(6), 657–674.
Google Scholar WorldCat

Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology*, *64*, 285–308.
Google Scholar WorldCat

Chomsky, N. (1995). *The minimalist program*. MIT Press.
Google Scholar Google Preview WorldCat COPAC

Chomsky, N. (2017). Language architecture and its import for evolution. *Neuroscience & Biobehavioral Reviews*, *81*, 295–300.
Google Scholar WorldCat

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5), 489–509.
Google Scholar WorldCat

Christiansen, M. H., & Chater, N. (2016). *Creating language: Integrating evolution, acquisition, and processing*. MIT Press.
Google Scholar Google Preview WorldCat COPAC

p. 898

Clark, E. V. (2009). *First language acquisition*. Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Clark, H. H. (1996). *Using language*. Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Clark, H. H. (2021). Anchoring utterances. *Topics in Cognitive Science*, *13*(2), 329–350.
Google Scholar WorldCat

Clark, H. H., & Schaefer, E. F. (1987). Concealing one's meaning from overhearers. *Journal of Memory and Language*, *26*(2), 209–225.
Google Scholar WorldCat

Cohen, A. S., & German, T. C. (2010). A reaction time advantage for calculating beliefs over public representations signals domain specificity for "theory of mind." *Cognition*, *115*(3), 417–425.
Google Scholar WorldCat

Colombo, M. (2013). Moving forward (and beyond) the modularity debate: A network perspective. *Philosophy of Science*, *80*(3), 356–377.
Google Scholar WorldCat

Cornejo, C., Cuadros, Z., Morales, R., & Paredes, J. (2017). Interpersonal coordination: Methods, achievements, and challenges. *Frontiers in Psychology*, *8*, 1685.
Google Scholar WorldCat

Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, 7, 114–133.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153.
Google Scholar WorldCat

Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in Psychology*, *2*, 355.
Google Scholar WorldCat

Dale, R., Fusaroli, R., Duran, N. D., & Richardson, D. C. (2013). The self-organization of human interaction. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 59, pp. 43–95). Academic Press.
Google Scholar Google Preview WorldCat COPAC

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: A comparative perspective. *Biological Reviews*, *89*(2), 284–301.
Google Scholar WorldCat

Dunning, D., & Balcetis, E. (2013). Wishful seeing: How preferences shape visual perception. *Current Directions in Psychological Science*, *22*(1), 33–37.
Google Scholar WorldCat

Falandays, J. B., Batzloff, B. J., Spevack, S. C., & Spivey, M. J. (2020). Interactionism in language: From neural networks to bodies to dyads. *Language, Cognition and Neuroscience*, *35*(5), 543–558.
Google Scholar WorldCat

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences*, *18*(3), 120–126.
Google Scholar WorldCat

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, *60*(6), 1497–1510.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*, e229.
Google Scholar WorldCat

Fitch, W. T., Huber, L., & Bugnyar, T. (2010). Social cognition and the evolution of language: Constructing cognitive phylogenies. *Neuron*, *65*(6), 795–814.
Google Scholar WorldCat

Fodor, J. A. (1983). *The modularity of mind*. MIT Press.
Google Scholar Google Preview WorldCat COPAC

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357–1392.
Google Scholar WorldCat

Freeman, J., Dale, R., & Farmer, T. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, *2*, 59.
Google Scholar WorldCat

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157.
Google Scholar WorldCat

Fusaroli, R., Bjørndahl, J. S., Roepstorff, A., & Tylén, K. (2016). A heart for interaction: Shared physiological dynamics and behavioral coordination in a collective, creative construction task. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(9), 1297–1310.
Google Scholar WorldCat

Gibbs, R. W. (1999). *Intentions in the experience of meaning*. Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Gibbs, R. W., & Bryant, G. A. (2008). Striving for optimal relevance when answering questions. *Cognition*, *106*(1), 345–369.
Google Scholar WorldCat

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427.
Google Scholar WorldCat

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

20(11), 818–829.
Google Scholar WorldCat

Goodwin, C. (2017). *Co-operative action*. Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Goodwin, C., & Heritage, J. (1990). Conversation analysis. *Annual Review of Anthropology*, *19*(1), 283–307.
Google Scholar WorldCat

Goodwin, M. H. (2017). Haptic sociality: the embodied interactive construction of intimacy through touch. In C. Meyer, J. Streek & J. S. Jordan (Eds.), *Intercorporeality: emerging socialities in interaction* (pp. 73–102). Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Graesser, A. C., Jeon, M., Yan, Y., & Cai, Z. (2007). Discourse cohesion in text and tutorial dialogue. *Information Design Journal*, *15*(3), 199–213.
Google Scholar WorldCat

Grice, H. P. 1975. Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 9: Pragmatics* (pp. 41–58). Academic Press.
Google Scholar Google Preview WorldCat COPAC

p. 899 Helmbrecht, J. (2003). Politeness distinctions in second person pronouns. In F. Lenz (Ed.), *Pragmatics and beyond new series* (pp. 185–202). John Benjamins.
Google Scholar Google Preview WorldCat COPAC

Heritage, J., & Clayman, S. (2011). *Talk in action: Interactions, identities, and institutions* (Vol. 44). Wiley.
Google Scholar Google Preview WorldCat COPAC

Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., Alemu, E., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., … Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*, *6*, 1545–1556.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*(3), 88–97.
Google Scholar WorldCat

Hurford, J. R. (2003). The language mosaic and its evolution. *Studies in the Evolution of Language*, *3*, 38–57.
Google Scholar WorldCat

Idemaru, K., Winter, B., & Brown, L. (2019). Cross-cultural multimodal politeness: The phonetics of Japanese deferential speech in comparison to Korean. *Intercultural Pragmatics*, *16*(5), 517–555.
Google Scholar WorldCat

Johnson, K. L., & Shiffrar, M. (Eds.). (2013). *People watching: Social, perceptual, and neurophysiological studies of body perception*. Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Kelly, S., Healey, M., Özyürek, A., & Holler, J. (2015). The processing of speech, gesture, and action during language comprehension. *Psychonomic Bulletin & Review*, *22*(2), 517–523.
Google Scholar WorldCat

Kingston, J., Levy, J., Rysling, A., & Staub, A. (2016). Eye movement evidence for an immediate Ganong effect. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(12), 1969–1988.
Google Scholar WorldCat

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

108–114.
Google Scholar WorldCat

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, *36*(8), 1404–1426.
Google Scholar WorldCat

Łuniewska, M., Haman, E., Armon-Lotem, S., Etenkowski, B., Southwood, F., Anđelković, D., Blom, E., Boerma, T., Chiat, S., Engel de Abreu, P., Gagarina, N., Gavarró, A., Håkansson, G., Hickey, T., Jensen de López, K., Marinis, T., Popović, M., Thordardottir, E., Blažienė, A., … Ünal-Logacev, Ö. (2016). Ratings of age of acquisition of 299 words across 25 languages: Is there a cross-linguistic order of words? *Behavior Research Methods*, *48*, 1154–1177.
Google Scholar WorldCat

Lupyan, G., Rahman, R. A., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences*, *24*(11), 930–944.
Google Scholar WorldCat

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, *110*(35), 14196–14201.
Google Scholar WorldCat

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*(2), 57–74.
Google Scholar WorldCat

Mesoudi, A., Laland, K. N., Boyd, R., Buchanan, B., Flynn, E., McCauley, R. N., Renn, J., Reyes-Garcia, V., Shennan, S., Stout, D. & Tennie, C. (2013). The cultural evolution of technology and science. In P. J. Richardson & M. H. Christiansen (Eds.), *Cultural evolution: Society, technology, language, and religion* (pp. 193–216). MIT Press.
Google Scholar Google Preview WorldCat COPAC

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*(5), 269–274.
Google Scholar WorldCat

Norris, D., McQueen, J. M., & Cutler, A. (2018). Commentary on "interaction in spoken word recognition models." *Frontiers in Psychology*, *9*, 1568.
Google Scholar WorldCat

Oller, D. K. (2000). *The emergence of the speech capacity*. Psychology Press.
Google Scholar Google Preview WorldCat COPAC

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences*, *110*(16), 6318–6323.
Google Scholar WorldCat

Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *The Quarterly Journal of Experimental Psychology*, *66*(11), 2092–2102.
Google Scholar WorldCat

Perner, J. (1991). *Understanding the representational mind*. MIT Press.
Google Scholar Google Preview WorldCat COPAC

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190.
Google Scholar WorldCat

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329–347.
Google Scholar WorldCat

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences*, *105*(3), 833–838.
Google Scholar WorldCat

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, *95*(2), 201–236.
Google Scholar WorldCat

Pouw, W., Paxton, A., Harrison, S. J., & Dixon, J. A. (2020). Acoustic information about upper limb movement in voicing. *Proceedings of the National Academy of Sciences*, *117*(21), 11364–11367.
Google Scholar WorldCat

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–365.
Google Scholar WorldCat

Quaglio, P., & Biber, D. (2006). The grammar of conversation. In B. Aarts & A. McMahon (Eds.), *The handbook of English linguistics* (pp. 692–723). Blackwell.
Google Scholar Google Preview WorldCat COPAC

M. A., & Turvey, M. T. (2001). The self-organizing dynamics of intentions and actions. *The American Journal of Psychology*, *114*(1), 160–169.
Google Scholar WorldCat

Ritwika, V. P. S., Pretzer, G. M., Mendoza, S., Shedd, C., Kello, C. T., Gopinathan, A., & Warlaumont, A. S. (2020). Exploratory dynamics of vocal foraging during infant–caregiver communication. *Scientific Reports*, *10*(1), 1–14.
Google Scholar WorldCat

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663–12668.
Google Scholar WorldCat

Scott-Phillips, T. (2015). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Macmillan International Higher Education.
Google Scholar Google Preview WorldCat COPAC

Shepherd, S. V., & Freiwald, W. A. (2018). Functional networks for social communication in the macaque monkey. *Neuron*, *99*(2), 413–420.
Google Scholar WorldCat

Shintel, H., & Keysar, B. (2009). Less is more: A minimalist account of joint action in communication. *Topics in Cognitive Science*, *1*(2), 260–273.
Google Scholar WorldCat

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*(2), 305–319.
Google Scholar WorldCat

Sperber, D. (2000). *Metarepresentations* in an evolutionary perspective. In D. Sperber (Ed.), *Metarepresentations* (pp. 117–138). Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition*. Blackwell Publishing.
Google Scholar Google Preview WorldCat COPAC

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, *65*(4), 235–241.
Google Scholar WorldCat

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*(7), 580–586.
Google Scholar WorldCat

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, *23*(2), 121–126.
Google Scholar WorldCat

Tollefsen, D., & Dale, R. (2012). Naturalizing joint action: A process-based approach. *Philosophical Psychology*, *25*(3), 385–407.
Google Scholar WorldCat

Tomasello, M. (1994). The question of chimpanzee culture. In R. W. Wrangham, W. C. McGrew, F. B. M. de Waal, & P. G. Heltne (Eds.), *Chimpanzee cultures* (pp. 301–377). Harvard University Press.
Google Scholar Google Preview WorldCat COPAC

Tomasello, M. (2009). *Constructing a language*. Harvard University Press.
Google Scholar Google Preview WorldCat COPAC

Tomasello, R., Kim, C., Dreyer, F. R., Grisoni, L., & Pulvermüller, F. (2019). Neurophysiological evidence for rapid processing of verbal and gestural information in understanding communicative actions. *Scientific Reports*, *9*, 16285.
Google Scholar WorldCat

Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580–591.
Google Scholar WorldCat

VanDam, M., Warlaumont, A. S., Bergelson, E., Cristia, A., Soderstrom, M., De Palma, P., & MacWhinney, B. (2016, May). HomeBank: An online repository of daylong child-centered audio recordings. In *Seminars in speech and language* (Vol. 37, No. 2, pp. 128–142). Thieme Medical.
Google Scholar Google Preview WorldCat COPAC

Vinson, D. W., Abney, D. H., Amso, D., Chemero, A., Cutting, J. E., Dale, R., & Spivey, M. (2016). Perception, as you make it. *Behavioral and Brain Sciences*, *39*, E260.
Google Scholar WorldCat

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324.
Google Scholar WorldCat

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*(3), 640–663.
Google Scholar WorldCat

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262.
Google Scholar WorldCat

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, *18*(5), 229–234.
Google Scholar WorldCat