Psychological Review

Cultural Evolutionary Pragmatics: Investigating the Codevelopment and Coevolution of Language and Social Cognition

Paula Rubio-Fernandez

Online First Publication, March 2, 2023. https://dx.doi.org/10.1037/rev0000423

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Review

ISSN: 0033-295X

https://doi.org/10.1037/rev0000423

Cultural Evolutionary Pragmatics: Investigating the Codevelopment and Coevolution of Language and Social Cognition

Paula Rubio-Fernandez
Department of Philosophy, University of Oslo

Language and social cognition come together in communication, but their relation has been intensely contested. Here, I argue that these two distinctively human abilities are connected in a positive feedback loop, whereby the development of one cognitive skill boosts the development of the other. More specifically, I hypothesize that language and social cognition codevelop in ontogeny and coevolve in diachrony through the acquisition, mature use, and cultural evolution of reference systems (e.g., demonstratives: "this" vs. "that"; articles: "a" vs. "the"; pronouns: "I" vs. "you"). I propose to study the connection between reference systems and communicative social cognition across three parallel timescales—language acquisition, language use, and language change, as a new research program for *cultural evolutionary pragmatics*. Within that framework, I discuss the coevolution of language and communicative social cognition as cognitive gadgets, and introduce a new methodological approach to study how universals and cross-linguistic differences in reference systems may result in different developmental pathways to human social cognition.

Keywords: Theory of Mind, pragmatics, reference, intersubjectivity, cognitive gadgets

Mastering communication takes more than mastering a language. Suppose that you and your partner are going to a concert: You may say "I forgot the tickets" to imply that you need to go back home. This simple example shows how, as speakers, we trust our listeners to read between the lines, and as listeners, we are willing to go beyond the literal to infer what the speaker intended to convey (Grice, 1975). Theoretical work on the nature of communication has long argued that communication requires social cognition, or a *Theory of Mind*: An ability to reason about mental states, such as beliefs, desires, and intentions (Levinson, 2006; Scott-Phillips, 2014; Sperber & Wilson, 1995; Tomasello, 2008). Yet, there is more to this ability than reading between the lines: By saying "the tickets," you would be signaling to your partner that these tickets are in your common ground (otherwise they would respond: "which tickets?"). The aim of this article is to explore the relation between language and social cognition by developing the hypothesis that the acquisition and regular use of reference systems (e.g., articles: "a" vs. "the"; demonstratives: "here" vs. "there," or pronouns: "I" vs. "you") train *communicative social cognition* (henceforth CSC): Those forms of social cognition that are required for successful communication, such as monitoring an interlocutor's visual attention in face-to-face interaction or their discourse memory during conversation.

For the last four decades, experimental research on Theory of Mind development has relied on *false-belief tasks*: a classic paradigm where a protagonist is mistaken about the location of an object and the child has to predict where the protagonist will look for the object, without defaulting to their own knowledge (Wimmer & Perner, 1983). False-belief tasks are widely considered as a litmus test for human social cognition, with more basic forms of Theory of Mind, such as monitoring others' attention or recognizing intentions, being treated as mere *precursors* (see Baron-Cohen, 1991, 1995; Whiten, 2000). This methodological bias, however, has given undue importance to false-belief understanding over other sociocognitive abilities. It is at least conceivable that an agent who could infer other agents' intentions but not their false beliefs would get further in their social world than an agent who could understand others' false beliefs, but not their intentions. Interestingly, according to the mainstream view, the agent without an understanding of false beliefs would not have a Theory of Mind.

Here, it is also important to remember that false-belief tasks were originally conceived as a reliable experimental paradigm for testing Theory of Mind, without false-belief attribution being necessarily more important than the attribution of other mental states. In their seminal article, Premack and Woodruff (1978) attributed a Theory of Mind to chimpanzees based on the observation that a 14-year-old chimpanzee was able to predict a human's solution to simple problems (e.g., obtaining out-of-reach fruit) by selecting the correct photograph from a set of alternatives. The authors' conclusion was criticized in three independent commentaries, arguing that the

Paula Rubio-Fernandez https://orcid.org/0000-0003-1622-0967

The author thanks the members of the Center for Language Evolution at the University of Edinburgh, the members of the Audience Design Group at the University of Rochester, and Damián Blasi, Daniel Harris, Richard Moore, Bart Geurts, and Marieke Woensdregt for many interesting discussions of the ideas presented in this article. Special thanks to Julián Jara-Ettinger for our own positive feedback loop. Paula Rubio-Fernandez received funding from FRIPRO Grant 275505 and Project Establishment Support (PES) from the Research Council of Norway (Norges Forskningsråd). The author has no conflict of interest to declare.

Correspondence concerning this article should be addressed to Paula Rubio-Fernandez, Department of Philosophy, University of Oslo, Georg Morgenstiernes hus, Blindernveien 31, 0313 Oslo, Norway. Email: paula.rubio-fernandez@ifikk.uio.no

chimpanzee might simply anticipate what it itself would do to solve the problem, without any mental-state attribution (Bennett, 1978; Dennett, 1978; Harman, 1978). Such a rich conclusion, these philosophers argued, would be more convincing if the chimpanzee had anticipated the human's acting differently from the way the chimpanzee would act in the same situation. This suggestion was followed up in developmental psychology a few years later with the birth of the false-belief task (Wimmer & Perner, 1983).

Given the field's narrow focus on false-belief tasks, research on the connection between language and social cognition has centered on children's acquisition of mental-state verbs (e.g., thinking, wanting and knowing; Bartsch & Wellman, 1995; de Villiers, 1999; Hacquard & Lidz, 2022). However, mental-state verbs pose high metarepresentational demands (e.g., consider the difference between "Sally knows the marble is in the basket" vs. "Sally thinks the marble is in the basket"), without being more important for successful communication (or social navigation) than other parts of language. As Evans et al. (2018a, 2018b; Bergqvist & Knuchel, 2019) have recently pointed out, human languages offer rich ways to track, compare, and engage the attentional and epistemic states of interlocutors through grammatical categories such as deixis and evidentiality—what they call the *grammar of engagement*.

Here, I will argue that, relative to traditional Theory of Mind research, the study of *intersubjectivity* (i.e., a speaker's assumptions about whether the listener shares their attention or memory) offers a window into earlier, more basic forms of CSC. In this view, monitoring our interlocutors' visual attention and memory for common ground could be seen as the building blocks for higher level forms of CSC, which are also trained by language (e.g., through the acquisition of mental-state verbs; see Taumoepeau & Ruffman, 2006, 2008).

Matsui et al.'s (2006) work on the acquisition of Japanese mental-state verbs is relevant to this proposal. Japanese encodes certainty and evidentiality in high-frequency sentence-final particles (e.g., "tte" marks hearsay), in addition to having low-frequency mental-state verbs (e.g., knowing vs. thinking). Matsui et al. (2006) showed that 3- to 6-year-old Japanese-speaking children understand the epistemic information encoded in sentence-final particles before they understand mental-state verbs. However, children's epistemic vocabulary correlated with their performance in standard false-belief tasks, whereas their understanding of sentence-final particles expressing the same meanings did not. Matsui et al. concluded that Japanese children's understanding of speakers' epistemic states as communicated by sentence-final particles paves the way for their later, fully representational understanding of belief.

The results of Matsui et al. (2006) highlight the importance of basic forms of CSC, above and beyond the acquisition of mental-state verbs. I will argue that these and other findings call for a paradigm shift away from the prevalent, methodologically skewed view of Theory of Mind as false-belief understanding, and toward a systematic investigation of those basic forms of social cognition that are equally fundamental to human communication: namely, monitoring an interlocutor's visual attention in face-to-face communication, and monitoring an interlocutor's memory for common ground.

Cultural Evolutionary Pragmatics

In the last two decades, evolutionary linguistics has made great progress in understanding the *cultural evolution of language* (i.e., the cultural emergence of linguistic structure through processes of learning and use; Christiansen & Kirby, 2003; Dor, 2015; Fitch, 2017; Planer & Sterelny, 2021; Smith & Kirby, 2008; Tamariz & Kirby, 2016). However, the received view is that CSC is innate and prior to the emergence of language. In fact, even those researchers who defend the cultural evolution view of human language and reject nativist accounts (e.g., the idea that humans are endowed with a Universal Grammar; Chomsky, 1965) nonetheless assume that the sociocognitive abilities involved in human communication are innate (e.g., Christiansen & Chater, 2016; Levinson, 2006; Scott-Phillips, 2014; Tomasello et al., 2005). Yet, none of these accounts have systematically explored the possibility that CSC and language may have coevolved (cf. Malle, 2002; Moore, 2021; Rubio-Fernandez, 2021a; Woensdregt et al., 2021) and may codevelop—although it is generally agreed that language must play a role in sociocognitive development (Wellman, 2014).

Seizing the maturity of this research area and the momentum of these research questions, I propose to investigate how language and CSC may have coevolved in diachrony and codevelop in ontogeny (i.e., coevolved during language change and codevelop during human maturation) through the acquisition, mature use, and cultural evolution of reference systems. This is an ambitious research program that requires addressing a fundamental question on the nature of human cognition across three parallel timescales: language acquisition, language use, and language change. This multiscale approach has been key to our understanding of the origins of human language as a product of cultural evolution (Christiansen & Chater, 2016; Dediu et al., 2013; Pleyer & Winters, 2014). Adopting these three parallel timescales to investigate the relation between language and CSC therefore has the potential to open a new research field: *cultural evolutionary pragmatics*.

The main aim of this article is to introduce a research program for cultural evolutionary pragmatics. To this end, I will first develop a new theory of the relation between language and CSC, and then outline an experimental program that is cross-linguistic and cross-cultural in scope and grounded in human lifespan development. The article is divided in two parts. In the first part, I develop the *positive feedback loop (PFL) hypothesis:* the idea that language and CSC are connected in a positive reinforcement cycle, and contrast it with earlier accounts. I start by reviewing previous developmental work showing a correlation between language and Theory of Mind, before developing the notion of intersubjectivity. I next introduce other key notions such as grammar, reference, and pragmatics, which lay the groundwork for the PFL hypothesis, and illustrate the workings of this reinforcement cycle with the case of common-ground management. To conclude the first part of the article, I draw parallels and differences between the PFL hypothesis and Heyes (2018) theory of *cognitive gadgets*.

In the second part of the article, I argue that the codevelopment and coevolution of language and CSC form *complex developmental pathways* (Smith, 2013). I introduce a new methodological approach to linguistic relativity, whereby the effect of language on CSC is investigated through their *joint use* in communication, rather than employing nonverbal tasks. I illustrate this novel approach by reviewing two recent cross-linguistic studies of demonstrative use, and conclude with a discussion of language change—the third timescale in cultural evolutionary pragmatics.

Part 1: A New Theory of the Relation Between Language and Social Cognition

Developmental Studies on Language and Theory of Mind

Recent cognitive neuroscience studies using functional magnetic resonance imaging (fMRI) observed that the brain region that supports Theory of Mind increases in functional specialization throughout childhood (ages 3–12 years; Richardson et al., 2018), and that deaf children with delayed exposure to American Sign Language show neural responses to Theory of Mind stories similar to those of young children with comparable linguistic experience, rather than to those of age-matched native-signing children with similar biological maturation (Richardson et al., 2020; for recent fMRI studies testing the relationship between the language and Theory of Mind brain networks with adults, see Jacoby & Fedorenko, 2020; Paunov et al., 2019, 2022). The results of these developmental studies therefore suggest that early linguistic experience may facilitate mentalizing through the development of a selective brain region for Theory of Mind (Richardson et al., 2020).

Behavioral studies have also documented an early connection between language and Theory of Mind in children's understanding of emotion, which closely parallels their language development (e.g., Beck et al., 2012; Cole et al., 2010; Cutting & Dunn, 1999; Ornaghi & Grazzani, 2013; Streubel et al., 2020). The early emergence of emotion understanding is thought to be supported by infants' discrimination of facial expressions of positive and negative emotions (Gori et al., 2021), which can also be labeled (e.g., "happy," "sad" or "angry"). Pons et al. (2004) have identified three successive phases of emotion understanding: between ages 3–5, children start to be able to recognize and name facial expressions of basic emotions; between 3 and 6 years, they understand that people have their own desires and beliefs, and begin to appreciate that mental states can trigger emotions; finally, from around 6 years onward, children achieve a reflective understanding of emotion, including their appreciation of hidden, mixed, and moral emotions.

Further supporting the connection between language and social cognition, numerous developmental studies have also shown a positive correlation between language and Theory of Mind (for a meta-analysis, see Milligan et al., 2007). These studies normally employ syntax and vocabulary scores as measures of linguistic ability, whereas Theory of Mind is assessed through false-belief tasks. Children do not pass standard false-belief tasks until around age 4 (Rakoczy, 2017), which is relatively late compared to the production of their first words at around 12 months (Kuhl, 2004; Schneider et al., 2015).

While providing robust evidence of the connection between language and Theory of Mind, correlational studies suffer from two methodological limitations. First, by relying on a late-emerging hallmark of Theory of Mind, they fail to establish *when* language and social cognition start to correlate in development. Second, by measuring these two cognitive capacities through separate tasks, correlational studies also fail to show *how* language and social cognition come to correlate. These two questions are at the heart of cultural evolutionary pragmatics.

A more specific theory of the connection between language and Theory of Mind was first put forward by de Villiers (1999, 2007), according to whom false-belief understanding emerges from children's mastery of sentential complement syntax (see also Moore, 2021). Analogous to a false-belief task, understanding "Sally thinks that the marble is in the basket" requires appreciating that the sentence may be true even though the marble is in the box (for supporting empirical evidence, see Boeg Thomsen et al., 2021; Lohmann & Tomasello, 2003).

The correlation between children's performance in false-belief tasks and their use of subordinate syntax is conceptually and developmentally sound, but possibly too narrow to sustain the interdependence of two foundational capacities in human cognition. In fact, syntax-based accounts of Theory of Mind development leave three fundamental questions unanswered, each related to a different timescale:

1. Regarding language acquisition, children do not pass standard false-belief tasks (or acquire complement clause syntax) until age 4 (Rakoczy, 2017), so *how does communicative social cognition develop up until that age?*

2. Regarding language use, once sentential complement syntax has been mastered, *how do mature speakers use their social cognition in everyday communication?*

3. Regarding language evolution, not all languages express mental states via subordinate clauses (Evans, 2006, 2021; Mithun, 1984), so *how did social cognition emerge across languages and cultures?*

I propose to address these open questions through the study of reference systems, which are (a) acquired early in development, (b) highly frequent in mature speech, and (c) used to mark intersubjectivity in all of the World's languages. I hypothesize that reference systems connect language and CSC and enable their coevolution in diachrony and codevelopment in ontogeny through a positive feedback loop, whereby the development of one skill boosts the development of the other (Rubio-Fernandez, 2021a, 2022).

Yet, before we turn to reference systems, a caveat is in order: The systematic study of the relation between language and CSC would make an important and timely contribution not only to cultural evolution research, but also to nativist accounts of both language and social cognition. After all, nativist accounts also need to explain how children *learn to use* their innate linguistic and sociocognitive abilities in communication. That is, any satisfactory account of human language and social cognition needs to not only identify which abilities may be innate (and which emerging), but also explain the development of these abilities as they are used in real-time social interaction during the lifespan. Therefore, investigating how human language and CSC codevelop in ontogeny is relevant to all researchers interested in these two foundational cognitive capacities, regardless of their stance on the nature/nurture debate.

Intersubjectivity as the Basis of CSC

Reference systems are formed by closed-class words (e.g., demonstratives, articles, or pronouns), which encode what some linguists call *procedural meanings*: Nonrepresentational information that is unavailable to consciousness and therefore implicit, but accessed automatically (Blakemore, 1987; see also Glanzberg, 2014; Harris, 2022). This explains why a competent user of English would understand that "I forgot the tickets" refers to some known-about

tickets, but would find it difficult to define the meaning of "the" (Gundel & Johnson, 2013). By contrast, conceptual meanings are conveyed by open-class words (such as nouns and verbs), which encode information that is representational and explicit, and therefore more accessible to introspection, but less automatic (e.g., the difference between thinking and knowing; see Grimshaw, 2005; Jackendoff, 1983, 2002; Jones, 2010; Pinker, 1989, 2007; Talmy, 2000).

The distinction between procedural versus conceptual meanings has been paralleled to that between implicit versus explicit Theory of Mind (Gundel et al., 2007; Matsui et al., 2006). In a similar vein, I argue that those forms of social cognition that are required for successful communication (what I call CSC) include not only the attribution of mental states, such as beliefs and desires (what is traditionally known as Theory of Mind), but also basic forms of social cognition, such as monitoring our interlocutors' attention and memory (what I call intersubjectivity). Other forms of intersubjectivity include the recognition of others' emotional states in their facial expression and tone of voice. However, even if language use requires both intersubjectivity and Theory of Mind, not all forms of social cognition have correlates in language.

For example, biological motion (i.e., the movement of other agents) conveys information that allows the identification of affective states and intentions (Federici et al., 2020), yet this human ability is not essential for language use and therefore falls outside the scope of CSC. Likewise, face recognition is a fundamental sociocognitive capacity that is sensitive to ingroup and outgroup differences (Prunty et al., 2022) but is not trained through language use. Therefore, while the connection between language and social cognition is both deep and complex, not all forms of social cognition are recruited and trained in communication.

One of the most basic forms of CSC is the ability to follow an interlocutor's gaze to resolve reference (Rubio-Fernandez et al., 2022), which allows interlocutors to establish *joint attention* (i.e., mutually sharing one's focus of attention with others; Tomasello, 2008). According to Tomasello (1999), the ability to engage in joint attention underpins the human capacity for cumulative culture, which forms the basis for our uniquely complex technologies, social systems, and languages. In the present account, joint attention is a form of intersubjectivity that is essential to communication, and which forms the basis for the interlocutors' common ground.

First introduced by the German philosopher Edmund Husserl at the start of the 20th century, the notion of intersubjectivity has been understood in various different ways across the social sciences (see, e.g., Duranti, 2010; Heasman & Gillespie, 2019; Pérez & Gomila, 2021; Zahavi & Overgaard, 2013). Trevarthen (1979, p. 347) concedes that "intersubjectivity is not a graceful word, but it does specify the linking of subjects who are active in transmitting their understanding to each other." Evans et al. (2018a) highlight the importance of intersubjectivity in the context of human language and communication: "Achieving intersubjectivity thus lies at the heart of how human communication systems evolved" (p. 3; see also Bergqvist & Knuchel, 2019; Verhagen, 2007).

Here, I will adopt the definition proposed by Stern (2005, p. 77), who understands intersubjectivity as "the ability to share in another's lived experience." I am partial to this formulation because it is compatible with both ongoing and past shared experiences between interlocutors. Thus, in situations of *copresence* (Clark & Marshall, 1981), interlocutors are mutually aware of sharing a physical environment, which affords them efficient choices in referential communication. For example, if something unusual was to happen in a situation of copresence, a vague message such as "Look at that!" may be informative enough—especially, if coupled with an ostensive head turn or a pointing gesture. On the other hand, if something unexpected happened while talking on the phone, the same surprised interlocutor would have to verbally describe what happened at their end of the line to the person on the other side, rather than relying on mutual salience (Jara-Ettinger & Rubio-Fernandez, 2022; Rubio-Fernandez, 2019).

Yet, in communication, sharing in past experiences (or the long-term common ground) is equally important when it comes to shaping reference. For instance, you and your next-door neighbor may safely refer to your janitor as "Fred," whereas if you were talking to someone outside your building, you would need to first explain that Fred is your janitor, before using his first name unambiguously. These decisions are also important when managing conversation in real time: Even if you had already introduced Fred in a conversation with someone visiting from another city, you would not be able to successfully refer back to Fred as "he" 10 min later—especially, if you and your visitor had started talking about other people in the meantime.

Seen this way, intersubjectivity is a basic form of social cognition that is based on sharing experiences, past or ongoing, and which is key to successful communication. In the present account, two forms of intersubjectivity are central to the study of cultural evolutionary pragmatics: monitoring an interlocutor's visual attention in face-to-face interction, and monitoring an interlocutor's memory, both discourse memory during conversation and long-term memory of past shared experiences. I further hypothesize that these two forms of intersubjectivity are *trained* by the early acquisition and highly frequent use of reference systems in all of the World's languages.

Key Ingredients: Grammar, Reference, Pragmatics, and Culture

The grammars of different languages place different constraints on their speakers. As Evans (2022) so eloquently put it in his paraphrase of Jakobson (2013): "Languages differ not so much in what you *can* say as in what you *must* say" (p. 71). Thus, speaking Dalabon requires distinguishing complex kinship relations, whereas speakers of Kayardilt must use case marking to distinguish different types of goals and intentions, while Andoke, Kogi or Marind demand that their speakers monitor their interlocutor's visual attention (Evans, 2022). Regarding intersubjectivity, different grammars have different engagement systems, including reference systems of diverse complexity (e.g., demonstratives that signal the location of a referent relative to the speaker's position, or redirect the listener's attention toward the intended referent) and extending to epistemic categories, such as evidentiality, which are not deictic in nature (Bergqvist & Knuchel, 2019; Evans et al., 2018a, 2018b).

The PFL hypothesis falls under the study of engagement, but it has a narrower scope, as it focuses on deictic systems as a lynchpin for the emergence and development of CSC in humans. In her discussion of the possible ways in which the emergence of language might have facilitated human social intelligence, Goody (1995) lists "reference by name" as the first and most obvious candidate, followed by the complex structures of grammar and syntax.

Hanks (2005) conceptualizes deictic practice as an *emergent social construal* (p. 198) and Agha (2007) characterizes reference as an *unavoidably social act* (p. 84), which is designed to succeed and thus achieve intersubjectivity (Haviland, 2007), but which might also fail and require repair (Heritage, 2007). In this view, an act of referring is successful to the extent that it allows speaker and listener to *mutually coordinate* on the intended referent (Agha, 2007). Stivers et al. (2007) thus claim that "reference entails a type of cooperation unique to humans" (p. 19; see also Enfield & Levinson, 2006).

In order to fully appreciate the necessarily social nature of referring, one must distinguish it from *denotation*: whereas referring is a highly contextualized activity that is dependent on the speaker's referential intention (e.g., which animal I mean when I talk about "my dog"), the denotation of an expression (e.g., the English word "dog") is its constant capacity to refer to the same type of entity across many acts of reference (for discussion, see Agha, 2007; Verhagen, 2015). Central to the PFL hypothesis (and its focus on reference systems) is the way in which deictic terms must be used and interpreted against the interlocutors' common ground. For example, the personal pronouns "I" and "you" refer to a different interlocutor every time two people take turns in a conversation, whereas their use of the third person pronouns "she" or "they" may refer to multiple individuals during the course of the same conversation (see Benveniste, 1971). Thus, using deictic expressions *presupposes* a common ground (Verhagen, 2015, p. 241).

Whereas the reference systems available in any one language are determined by its grammar, the processes of *referential choice* (during language production) and *reference assignment* (during language comprehension) are pragmatic processes (Carston, 2002). In other words, formulating and understanding referential expressions are instances of contextualized language use, including its social context. Thus, referring to one's father as "Dad," "my old man," "your grandpa" or "John" is a pragmatic choice that one makes for a specific audience in a specific context, just as understanding who these expressions refer to would depend on the speaker and the context of utterance. This is an important observation that places the PFL hypothesis at the level of pragmatics, not grammar—although this hypothesis is both constrained and supported by grammar (see Barth et al., 2021). In this view, the acquisition and regular use of reference systems trains CSC through *pragmatic reasoning*: speakers must select referential expressions that enable their recipients to identify the intended referent (a process known as *audience design*), while listeners must reason about the speaker's referential intention when interpreting these expressions (Frank & Goodman, 2012; Grice, 1975; Jara-Ettinger & Rubio-Fernandez, 2021, 2022). Therefore, empirically testing the PFL hypothesis (or making progress in our understanding of cultural evolutionary pragmatics, more generally) would require the expertise and collaboration not only of researchers in linguistics, typology, and social cognition, but also in experimental pragmatics, psycholinguistics, developmental psychology, and cognitive science and neuroscience.

This article will focus on cross-linguistic variability in reference systems and its potential effects on CSC, yet cross-cultural diversity is equally important in that respect (Enfield & Levinson, 2006; Robbins & Rumsey, 2008; Stivers et al., 2007). For example, Sacks and Schegloff (2007) observed two organizing principles for determining how person reference should be formulated: First, an overriding preference for *recognitionals* (i.e., referential expressions that allow the listener to identify the intended person) and second, a preference for minimized reference. Whereas these two preferences are concurrently satisfied in the use of proper names (e.g., "Nana" or "Mrs. Dalloway"), cross-cultural research has revealed clear differences in the degree to which the use of proper names and kinship terms is appropriate in conversation (Enfield, 2006; Garde, 2008, 2011; Haviland, 2007). More generally relevant to the role of social cognition in communication, a widespread assertion in the societies of the Pacific is that it is impossible, or at least extremely difficult, to know what other people think or feel. This idea, known as the *doctrine of the opacity of other minds*, has often been raised as a challenge to Western-centric views of communication and social interaction based on intention recognition and pragmatic reasoning (for discussion, see Robbins & Rumsey, 2008).

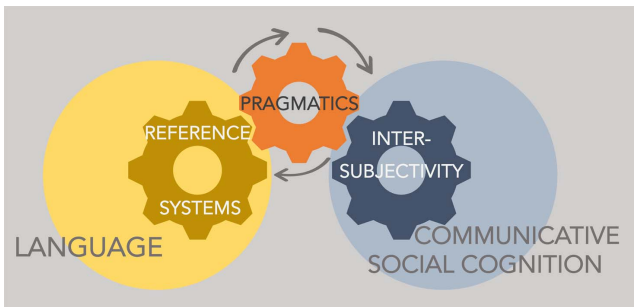
These and many other studies therefore confirm that cultural anthropologists would also need to collaborate with our long list of experts in order to make definite progress in cultural evolutionary pragmatics. After all, if it takes a village to raise a child, it probably takes just as many social and cognitive scientists to understand the origins and development of that child's mind.

The Positive Feedback Loop Between Language and CSC

Biologists could also help explain the mechanics of a positive feedback loop, which occurs when something needs to happen quickly in a biological system. In child birth, for example, when labor begins, the baby's head pressure on the cervix stimulates receptor cells that send a chemical signal to the brain, allowing the release of oxytocin, which in turn triggers contractions that result in more oxytocin release until the baby is born. Feedback processes can also be found in sociology, meteorology, economics, and chemistry, to list just a few. Systems in these and other areas operate on mechanisms with inputs and outputs, where each is caused by and causes a certain effect. A positive feedback loop is therefore a self-regulating mechanism wherein a positive output amplifies the system. Or in diagram terms: *A produces more of B, which in turn produces more of A* (Keesing, 1981).

According to the PFL hypothesis, the acquisition and regular use of reference systems trains CSC via pragmatic reasoning, which in turn boosts referential communication (see Figure 1). This hypothesis does not exhaust the complex interconnectivity between language and CSC, which extends to other engagement systems

Figure 1
Diagram of the Positive Feedback Loop Hypothesis



Note. See the online article for the color version of this figure.

(e.g., evidentials), as well as syntactic constructions (e.g., subordinate clauses) and mentalistic vocabulary (e.g., mental-state verbs and emotion adjectives). However, unlike these other connections (which might form other positive feedback loops with their own developmental timelines), reference systems emerge very early in both diachrony and ontogeny, and are employed with high frequency in all of the World's languages. Therefore, the coevolution and codevelopment of language and CSC is likely to have started with the emergence of reference—including pointing gestures (Kita, 2003; Özyürek, 2014).

In ontogeny, the human ability for shared reference is firmly established from around 12 months, when infants being to point (Liszkowski, 2018). This early ability to engage in joint attention with others is fundamental to children's communicative development (Tomasello, 2008) and is soon used to manage common ground in referential communication. Developmental studies have shown that when 14-month-olds share some objects with an adult, they are later able to resolve an ambiguous request by distinguishing which objects are familiar to the adult, and which one is new (Moll & Tomasello, 2007; Moll et al., 2007, 2008; Tomasello & Haberl, 2003; see also O'Neill, 1996). Interestingly, if the adult's request does not convey excitement about something new, infants have a general tendency to select those objects that are mutually familiar through shared experience (Liebal et al., 2009, 2011; Saylor & Ganea, 2007). Developmental studies therefore confirm that joint attention shapes infants' memory of shared referents, which in turn allows them to distinguish what is new and familiar to others (Moll & Kadipasaoglu, 2013).

The acquisition and regular use of reference systems builds on the basic distinction between what is new and familiar to our interlocutors. Thus, through the systematic demands of their grammars, children come to regularly mark this distinction in their language use and register it in their language comprehension (e.g., via the use of articles in English, as in the opening example: "I forgot the tickets"). The obligatoriness and high frequency of these markers ultimately result in the automatization of common-ground management in mature speakers.

Above and beyond the specific demands posed by our grammars, two other experiential factors are likely to play a key role in the positive reinforcement cycle between language and CSC. One is miscommunication, which normally requires repair in both children and adult conversation (Clark, 2020; Dingemanse et al., 2015; Forrester & Cherington, 2009). In instances of reference failure, interlocutors often try to reformulate their referential expressions to achieve intersubjectivity (Haviland, 2007), offering a window into the pragmatic reasoning that underlies referential choice (Carmiol et al., 2018; Clark & Krych, 2004; Clark & Wilkes-Gibbs, 1986; Matthews et al., 2007). Miscommunication and repair can therefore reinforce the positive connection between reference systems and CSC.

Second and more fundamentally, perhaps, is speakers' extensive experience as listeners, through which our language comprehension comes to inform our language production (Ferreira, 2019; Rubio-Fernandez, 2021b). Recent cross-linguistic studies comparing the production and real-time comprehension of modified descriptions (e.g., "the red cup" vs. "the plastic cup") have revealed that speakers use redundant (or noncontrastive) modification when it facilitates their listener's visual search for a referent given their adjective position (i.e., whether it is prenominal or postnominal; Rubio-Fernandez, 2021b; Rubio-Fernandez & Jara-Ettinger, 2021; Rubio-Fernandez et al., 2021; for computational models and further empirical evidence, see Jara-Ettinger & Rubio-Fernandez, 2022). Extensive experience with referential communication may also result in the development of *heuristics*, as when perceptual contrast triggers the use of redundant modification (Ferreira, 2019; Long et al., 2021). Therefore, experimental studies and computational models suggest that our experience as listeners informs the choices we make as speakers in referential communication. The extent to which our listener experience facilitates our audience design (i.e., the tailoring of referential expressions for our interlocutors; Ferreira, 2019) is an important empirical question that would help us understand how language use promotes perspective taking.

The Automatization of Common-Ground Management

According to Apperly (2010), everyday mindreading processes undergo *downward modularization* through extensive social interaction, comparable to the practice effects observed with other complex tasks that can eventually be automatized (e.g., driving a car or playing a musical instrument). Thus, learning to read minds may require reasoning about the thoughts of others, but practice can take this out of the equation (at least in everyday mindreading processes), enabling very fast performance by parsing others' behavior in an expert manner (see also Perner, 2010). In the realm of language and communication, Goody (1995) argues that automatization frees many complex processes from awareness, such as the order in which we string our words, or the motor control of speaking. Interestingly, these and other production processes might reach consciousness in the early stages of learning a second language as an adult. In line with these views, I hypothesize that the acquisition and regular use of reference systems bootstraps infants' early sociocognitive skills (Moll & Tomasello, 2007; Moll et al., 2007, 2008; Tomasello & Haberl, 2003), resulting in the downward modularization of common-ground management over time.

A recent self-paced reading study investigated the automatization of common-ground management, showing that adults derived automatic belief inferences when their common-ground expectations were violated in a dialog (Rubio-Fernández et al., 2019). Participants slowed down their reading when a stranger in a vignette made a passing comment that suggested private knowledge (e.g., "You are on a bus when a stranger tells you: 'That woman over there reminds me of your ex-wife'"). However, when the stranger commented on their own personal life (e.g., "That woman over there reminds me of my ex-wife") or a friend made either type of comment, participants did not slow down their reading. Note that in the stranger condition, a significant difference emerged in the processing of a possessive determiner (e.g., "your ex-wife" vs. "my ex-wife"), confirming that reference systems are a lynchpin for the automatization of common-ground management in adult communication.

Language acquisition studies, on the other hand, have revealed that reference systems are acquired early (especially demonstratives and pointing gestures; Clark, 1978), but common-ground management has a protracted development. Cross-linguistic studies using narrative-elicitation tasks have revealed that children as young as 3 years use pronouns correctly to maintain reference, using appropriate forms for familiar characters by the age of 4, whereas new characters are not introduced appropriately before age 7, and character reintroduction is not mastered until age 10 (Hickmann & Hendriks, 1999; Hickmann et al., 2015). Wong and Johnston (2004) have argued that the order of acquisition of these three

discourse functions (i.e., reference maintenance > introduction > reintroduction; Bamberg, 1986, 1987) results from their different sociocognitive demands, as speakers must maintain a mental model of the ongoing discourse in order to keep track of the listener's knowledge (Levelt, 1989). In their view, reference introduction and maintenance require information about the listener's knowledge to be updated, but reintroduction makes a further demand: that the listener's focus of attention be monitored. Thus, young children may be familiar with the appropriate linguistic forms for reintroduction (e.g., definite articles in English), but their failure to track the listener's attentional focus leads to inappropriate referencing (e.g., by using an ambiguous pronoun; Wong & Johnston, 2004).

Cross-linguistic studies on children's acquisition of articles have reported two types of errors (De Cat, 2011, 2013; Fuchs et al., 2021): *discourse integration errors* (i.e., using an indefinite article to refer to a familiar character; e.g., "A dog is sleeping," when the dog had already been introduced) and *egocentric errors* (i.e., wrongly assuming common ground and using definite articles for new characters; e.g., "The dog is sleeping," when the dog had not yet been mentioned). While there is considerable variability in the error rates reported in different studies (probably due to methodological differences), young children tend to make more egocentric errors than discourse integration errors (Power & Dal Martello, 1986; Warden, 1976; cf. Fuchs et al., 2021). These acquisition errors stand in stark contrast with adults' automatic distinction between shared and private referents in their selective use of articles with different interlocutors (see Bard & Aylett, 2005; Bard et al., 2000; Clark & Wilkes-Gibbs, 1986; Hupet & Chantraine, 1992). Thus, a mature language user would readily understand that "We bought the house" refers to the particular house that the speaker had recently showed them, but would formulate the same message as "They bought a house" if their interlocutor was not familiar with the house in question.

Supporting the PFL hypothesis, definite articles have been found to have a positive effect on memory recall (Bransford & Franks, 1971; de Villiers, 1974; for related evidence with pronouns, see Lesgold, 1972). Adult participants in Irwin et al. (1982) recognized English words faster if they had seen these words in a previous text than if it was the first time they were presented with a target. Crucially, this priming effect was greater if the target noun was marked with a definite article (e.g., "the kite") than an indefinite article ("a kite"). Other psycholinguistic studies have shown that definite expressions are understood faster than indefinite expressions (Murphy, 1984), especially when the same discourse referent (or *antecedent*) was recently mentioned (Clark & Sengul, 1979; Haviland & Clark, 1974). The results of these early studies suggest that definite articles have a facilitatory effect on memory search (for theoretical and computational models of discourse memory, see Garrod & Sanford, 1982; Gerrig & O'Brien, 2005; Pyke et al., 2007a, 2007b).

Work with amnesic patients further confirms the connection between articles and declarative memory (Duff et al., 2011). Amnesic patients playing a referential communication game successfully established common ground in their recurrent use of unique labels for tangram figures (e.g., calling the same abstract shape "Viking ship" across trials). However, these patients revealed a memory impairment in their inconsistent use of articles: Even when they had referred to a tangram figure several times, they continued to use indefinite articles (e.g., "a Viking ship"), incorrectly signaling that this was a novel referent. Neurotypical controls, on the other hand, made correct use of the definite article to signal referents in common ground (e.g., "The chimney" in a subsequent mention), while reserving the indefinite article for novel tangrams (for a review of studies of memory for common ground, see Brown-Schmidt & Duff, 2016).

Whereas not all languages have articles, marking which information is new and which information is familiar to our interlocutors is a universal function of all languages (Evans & Levinson, 2009; Ford & Fox, 1996; Givón, 1992, 2001). Therefore, speakers of all languages must use the grammatical devices at their disposal (e.g., articles, case marking, word order, ergative or split ergative alignment; see Evers, 2020) to manage their common ground with their interlocutors, all while monitoring their shared discourse and long-term memory. Since reference systems are highly frequent words in all of the World's languages, and distinguishing new and familiar information is a universal function of all languages, I predict that similar processes of downward modularization should be observed in common-ground management across all languages. However, it is an open empirical question whether and how the different grammatical means whereby speakers of different languages mark common ground may affect their CSC and its automatization (for discussion, see Evans, 2022).

For instance, the above studies confirm the connection between adults' use of articles and their memory for common ground, yet these investigations were limited to English (see Blasi et al., 2022). In fact, no psycholinguistic study to date has compared the effect of using different types of definiteness markers in the processing and recall of familiar referents across different languages. Mandinka, for example, uses demonstratives, instead of definite articles, to signal familiarity (Creissels, 2020), whereas bare nouns in Hindi can have a definite interpretation (Dayal, 2018). The cross-linguistic question therefore remains: would demonstrative determiners in Mandinka and bare nouns in Hindi reveal the same facilitatory memory effects (in both sentence processing and memory recall) than the definite article in English? (e.g., Clark & Sengul, 1979; Haviland & Clark, 1974; Irwin et al., 1982; Murphy, 1984). Addressing this important question (with these and other languages) should be one of the goals of cultural evolutionary pragmatics, as it would deepen our understanding of the downward modularization of common-ground management through the highly frequent use of different reference systems in everyday communication.

The Cultural Evolution of Cognitive Gadgets

Heyes (2018, 2019a) has recently argued that uniquely human cognitive mechanisms such as imitation, Theory of Mind and language are cognitive gadgets that have been shaped by cultural rather than genetic evolution:

> In cultural learning, what the learner learns depends on what the model or teacher knows. When cognitive gadgets are culturally learned, input from the sender enables the receiver's cognitive system to reconstruct the mental software that generated the sender's modeling or teaching behavior. Thus, the information that builds a cognitive gadget from old parts (i.e., psychological processes shared with other animals) comes from other people who have already acquired the mechanism. It is inherited like money rather than like eye color, through cultural learning rather than DNA. (Heyes, 2020, p. 400)

In the case of Theory of Mind, Heyes (2018; see also Heyes & Frith, 2014) argues that learning to "read minds" is a form of cultural

inheritance in which adults instruct children about the mind *by talking about other people's minds*. Importantly, parents scaffold the development of mindreading by first introducing easy-to-grasp mental states, such as desires and emotions, before they mention harder mental states, such as beliefs and knowledge. Longitudinal studies have shown that the frequency of developmentally apt references to mental states predicts children's Theory of Mind development (Taumoepeau & Ruffman, 2006, 2008). A product of cultural learning, Theory of Mind is a cognitive gadget that emerges from expert mindreaders communicating mental state concepts to Theory of Mind novices, who will in turn become the next generation of mindreading experts.

Cultural evolutionary pragmatics—the empirical study of the coevolution of language and CSC through processes of learning and use, is compatible with Heyes' view that genetic evolution has made subtle changes to the human mind such that human children can develop cognitive gadgets through cultural learning (Heyes, 2020). In this view, cognitive gadgets are adaptive because they are primarily shaped by culture, rather than nature or nurture (Heyes, 2019a). However, while compatible with Heyes' cultural evolutionary account of cognitive gadgets, the PFL hypothesis assumes a deeper, more fundamental interdependence between language and CSC (see also Bergqvist & Knuchel, 2019; Evans et al., 2018a, 2018b).

Ultimately, it is at the discretion of every parent (and every culture) how much they should talk about other people's minds with their children (see Robbins & Rumsey, 2008). This variability should indeed result in individual differences in mindreading skills (for discussion, see Heyes, 2019b). However, it is not at the discretion of every parent which grammatical rules to obey or flout when talking to their children. As a native speaker of English, every parent communicating with their child in that language will normally use indefinite articles to introduce new referents (e.g., "Today I saw *a fox* in the garden") and pronouns and definite articles when talking about familiar referents ("but when *it* saw me, *the fox* ran away!"). This cultural convention (or equivalent ones in all of the World's languages; Evans & Levinson, 2009) trains both parents and children in keeping track of common ground and ultimately automatizing basic mindreading processes through communication. In this view, language is more than a vehicle for "talking about the mind": The very processes of acquiring and using language recruit sociocognitive mechanisms that are fundamental for successful communication.

Part 2: Testing the PFL Hypothesis

CSC Develops Through Complex Pathways

Experimental research with infants has long established developmental connections between sensorimotor development, visual object recognition and word learning (Smith et al., 2018; Thelen & Smith, 1994, 2007), supporting a "pathways approach" to human development and evolution (Smith, 2013). For example, the ability to sit steadily and manipulate objects is part of the developmental pathway leading to visual object recognition, which is in turn fundamental to word learning. Smith (2013) identifies two ways in which developmental pathways to specific outcomes are complex: They are *multicausal* (i.e., each change is dependent on multiple causes) and also *degenerate* (in the biological sense: There is more than one route to the same functional end).

In this view, referential systems worldwide would help speakers develop their intersubjective skills, since speakers of all languages must be able to direct their interlocutor's attention to an intended referent, or mark new and familiar information for their interlocutor's benefit (Evans & Levinson, 2009; Evers, 2020; Ford & Fox, 1996; Givón, 1992, 2001; Kita, 2003; Özyürek, 2014). However, cultural evolutionary pragmatics should investigate not only those referential functions that are universal, but also how cross-linguistic differences determine which intersubjective skills are ultimately automatized by acquiring and regularly using each reference system. It is only through this dual approach focusing on both universals and cross-linguistic differences that we will come to understand how exactly language trains CSC across all languages and cultures.

Here, it is also important to bear in mind that by conceptualizing the codevelopment of language and CSC as complex developmental pathways, it follows that different individuals and different speech communities may develop comparable linguistic and sociocognitive abilities through different routes—including their having different individual experiences and acquiring different languages. Thus, signaling familiarity through a definite article or an accusative case marker (as speakers of English and Turkish do, for example) may result in equivalent processes of common ground marking. However, other cross-linguistic differences in reference systems may result in the automatization of different intersubjective processes in different speech communities.

A New Methodological Approach to Linguistic Relativity

The positive effect of language on CSC is a form of what Slobin (1996) called *thinking for speaking*: Speakers automatically monitor those features of the environment that are encoded in their grammars (Boroditsky, 2011; Papafragou et al., 2008; Wolff & Holmes, 2011). For example, depending on the language, demonstratives may indicate not only the distance of a referent from the speaker's position (as in "here" vs. "there" in English), but also a referent's altitude, familiarity, position, reachability, or visibility, from the perspective of the speaker, the listener, or both (Levinson, 2018). Therefore, depending on the relational values and perspectives encoded in a given demonstrative system, speakers of two different languages may come to automatically monitor whether a referent is close to the speaker, or visible to the listener (Rubio-Fernandez, 2021a, 2022).

Of the many formulations of the linguistic relativity hypothesis that have been proposed over the years (see Enfield, 2015), thinking for speaking has been the least controversial one (Wolff & Holmes, 2011). However, defendants of nativist accounts of conceptual structure have often downplayed the importance of thinking for speaking on the grounds that "any language-driven effects on cognition should be malleable and ephemeral in nature" (Ünal & Papafragou, 2020, 125) because "these language-driven effects are temporary and task-dependent (…) without modifying the underlying conceptual representations" (Ünal & Papafragou, 2018, 165).

The scope of cultural evolutionary pragmatics—as conceived here, falls outside this long-standing debate for two reasons: First, empirical research on linguistic relativity has focused on cross-linguistic differences that may affect thought, without regard for the effect of linguistic universals on thought. Second, nativists have called for the effects of language on thought to be measured through

nonverbal tasks, dismissing comparable effects observed with verbal tasks (see, e.g., Ünal & Papafragou, 2018, 2020). However, investigating how language and CSC may have coevolved and codevelop requires focusing both on universals and cross-linguistic differences, and investigating, precisely, how language and social cognition are jointly used in communication. Notice that this is the case even if one assumes that the human conceptual repertoire is innate and universal, since such theories must also explain how these innate concepts are deployed in interaction and may mature through experience (see, e.g., Heiphetz et al., 2013).

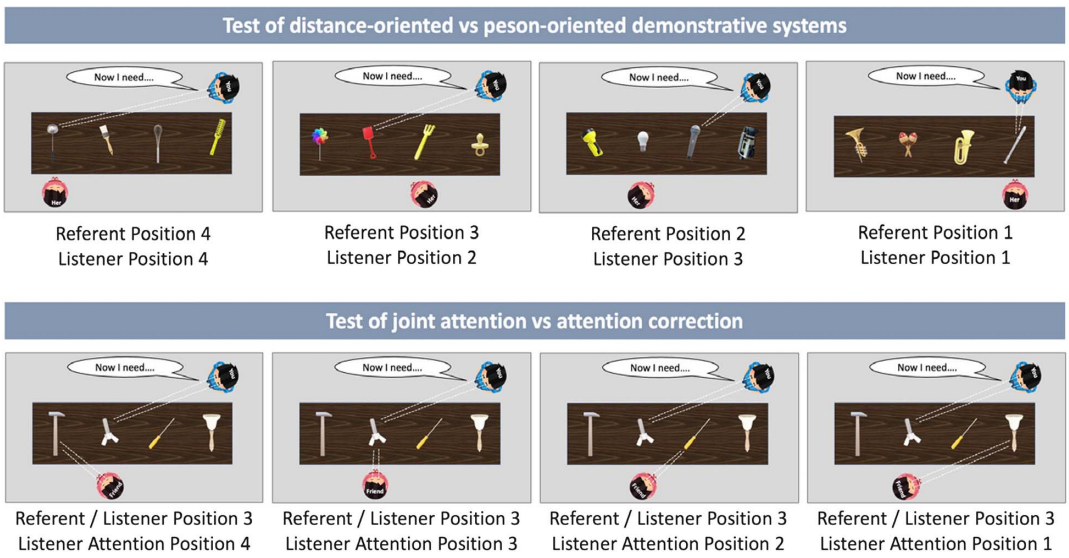
To better understand the positive effect of language on social cognition, I propose to investigate how these two cognitive capacities are jointly used in verbal interaction, focusing on those specific forms of social cognition that are essential for successful communication. This methodological approach differs markedly from previous studies of the effect of language on thought, which traditionally employed nonverbal tasks with speakers of different languages (for a review, see Enfield, 2015). The artificial nature of those tasks, however, has been argued to undermine the depth and even the validity of the effects of language on thought that these tasks revealed (Gentner & Goldin-Meadow, 2003; cf. Lupyan et al., 2020; Majid, 2021). Because cultural evolutionary pragmatics aims to explain how language trains social cognition through communicative interaction, establishing correlations between language assessments and false-belief tasks (Milligan et al., 2007), or measuring the effect of language on nonverbal tests of social cognition (Ünal & Papafragou, 2020) will not get to the bottom of this question. Instead, empirical studies need to use pragmatic tasks to investigate the effect of universals and cross-linguistic differences in reference systems on CSC. Then, once we better understand how language trains social cognition in communication, future studies should investigate the effect of language on social cognition outside of communication.

Universals and Cross-Linguistic Differences in Demonstrative Choice

To illustrate the importance of universals and cross-linguistic differences in reference systems for the study of cultural evolutionary pragmatics, I will review the results of a recent online study investigating demonstrative choice in various languages (Rubio-Fernandez, 2022; see also Woensdregt et al., 2022). Participants in four experiments were asked to adopt the role of the speaker in a visual scene (see Figure 2) and ask a friend to pass them one of four objects along a table. To complete the speaker's request (shown inside a speech bubble in the language of test; e.g., "Now I need …" in English), participants were given a choice of two or three demonstrative pronouns, depending on the language (e.g., "this one"/"that one" in English, or "kore"/"sore"/"are" in Japanese).

In the first three experiments, the speaker in the visual scene was always on the top right corner of the table, whereas the positions of the referent and the listener varied parametrically across trials (for sample displays, see top row in Figure 2). This setup was used to test of the typological distinction between *distance-oriented* versus *person-oriented demonstrative systems* (Diessel, 2013). In person-oriented

Figure 2
Sample Displays From Demonstrative-Choice Task



Note. Sample displays from "Demonstrative Systems: From Linguistic Typology to Social Cognition," by P. Rubio-Fernandez, 2022, *Cognitive Psychology*, *139*, Article 101519 (https://doi.org/10.1016/j.cogpsych.2022.101519). CC BY 4.0. The four scenes in the top row were used to test the typological distinction between distance-oriented and person-oriented demonstrative systems with native speakers of Catalan, English, Japanese, Spanish (monolingual and Catalan bilingual samples), and Turkish. The four scenes in the bottom row were used to compare the grammatical strategies that native speakers of Japanese, Spanish, and Turkish employed for attention correction (i.e., flexible use of the proximal and distal demonstratives to redirect the listener's attention, or systematic preference for the medial demonstrative to signal misaligned perspectives). See also Woensdregt et al. (2022) for computational models of demonstrative systems, and data from English, Italian, Portuguese, and Spanish speakers. See the online article for the color version of this figure.

systems, proximal demonstratives indicate proximity to the speaker, whereas the medial forms indicate distance from the speaker but proximity to the listener, and the distal forms indicate distance from both interlocutors. The results confirmed that Japanese, and Spanish have person-oriented systems, with participants selecting the distal demonstrative when the referent was at the other end of the table, far away from both speaker and listener; but as the listener got closer to the referent in other trials, participants showed a preference for the medial form. By contrast, the results confirmed that Catalan, English, and Turkish have distance-oriented demonstrative systems, where distance is always established from the speaker's position.

It follows from the PFL hypothesis that children learning Japanese, Spanish, and other languages with person-oriented systems (e.g., Portuguese; see Woensdregt et al., 2022) should eventually come to automatically monitor their listener's spatial location in order to accurately use their demonstratives, whereas this form of intersubjectivity should not be automatically deployed when native speakers of Catalan, English, Turkish, or other languages with distance-oriented systems (e.g., Italian; see Woensdregt et al., 2022) use demonstratives (for a recent study of Ticuna, see Skilton, 2022). Interestingly, when Catalan–Spanish bilinguals were tested in Spanish, they revealed sensitivity to the listener position only in their use of the distal form "aquel" (gloss: *far from me and you*), but not in their use of the medial "ese" (gloss: *far from me*), which was different from the pattern observed with monolingual Spanish speakers (gloss: *far from me and close to you*). These results were interpreted as a transfer effect from Catalan (which has a 2-way distance-oriented system), confirming that speaking two languages with different demonstrative systems can attenuate speakers' sensitivity to listener position in the language with a person-oriented system.

Also known as *directives*, demonstratives' primary function is to orient the listener's attention toward an element in the speech situation, often accompanied by a pointing gesture (Diessel, 2006; Kita, 2003). Since all languages have demonstratives (Levinson, 2018), the use of these forms as directives offers an opportunity to investigate how a linguistic universal may recruit a fundamental form of intersubjectivity; namely, monitoring the listener's visual attention. This was tested in the last experiment of the study with native speakers of Japanese, Spanish, and Turkish (Rubio-Fernandez, 2022; Experiment 4). Results revealed that native speakers of all three languages are sensitive to listener attention in their choice of demonstratives, but did so through two different grammatical strategies.

Japanese and Spanish speakers used the proximal and distal demonstratives flexibly to reorient the listener toward the intended referent (for sample displays, see bottom row in Figure 2). Thus, when the listener was looking closer from the target object, participants showed a preference for the distal form (gloss: *Look over there!*), whereas when she was looking further from the target, they used the proximal form more often (gloss: *Look over here!*). By contrast, Turkish speakers showed a preference for the medial demonstrative pronoun "şuna" in all trials in which the speaker and listener perspectives were misaligned, irrespective of the direction in which the listener had to turn to find the referent. In other words, Turkish has *lexicalized* the medial demonstrative "şu" for attention correction (Özyürek, 1998; for similar typological analyses of Yucatec Mayan, see Bohnemeyer, 2018; Hanks, 2009).

Since all demonstrative systems have a directive function (Diessel, 2007; Levinson, 2018), it follows from the PFL hypothesis that speakers of all languages should show sensitivity to the listener's focus of attention in their demonstrative use. The different grammatical strategies employed by Japanese, Spanish, and Turkish speakers illustrate how the acquisition and regular use of demonstratives may train speakers' monitoring of their listeners' attention across languages and cultures.

More generally, the online experiments in Rubio-Fernandez (2022) and Woensdregt et al. (2022) offer a new methodological approach to study the effect of language on CSC; namely, through pragmatic tasks that can reveal the demands that different reference systems pose on CSC, with some demands being universal, while others are language specific. In the case of demonstratives, cross-linguistic differences in speakers' monitoring of the listener's spatial location (which differentiates person-oriented vs. distance-oriented systems) and the universal use of demonstratives to redirect the listener's attention toward an intended referent are relevant to the study of cultural evolutionary pragmatics in so far as both reveal the interdependence of language and CSC.

A Diachronic View of Common Ground: From Demonstratives to Definite Articles

I started this article by arguing that the relation between language and social cognition needs to be investigated across three parallel timescales: language acquisition, language use and language change (Rubio-Fernandez, 2021a). However, the ensuing discussion focused on the first two of these timescales, leaving language change until last. In addition, I made a case for the study of two basic forms of intersubjectivity that are essential to human communication: namely, speakers' monitoring of listeners' visual attention in face-to-face interaction, and speakers' monitoring of listeners' memory for common ground in conversation. In what follows, I will show how adopting a diachronic perspective offers a connection between those grammatical forms used in all languages to direct an interlocutor's attention (i.e., demonstratives) and those forms used in some languages to signal common ground (i.e., definite articles). In other words, language change offers a grammatical link between monitoring an interlocutor's attention and memory.

Central to language change is the notion of *grammaticalization*: the process whereby content words develop into grammatical markers (Comrie, 1989; Diessel, 2007; Hopper & Traugott, 2003). These processes tend to have a common source and follow *universal pathways*: for example, demonstratives tend to evolve into definite articles, but not the other way around (Greenberg, 1978; Lyons, 1999). In their most basic exophoric function, demonstratives have the same role as a pointing gesture: to indicate the location of a physical referent (e.g., "Look at that house!"). These forms often develop into anaphoric demonstratives, directing the listener's attention to a discourse referent (e.g., "That was a good year"). At a later stage, anaphoric demonstratives provide a common historical source for definite articles (e.g., "We bought the house"; Diessel, 2006, 2007).

I have recently put forward a diachronic analysis of common ground whereby this pathway of language change marks a three-step expansion of the speakers' working model of common ground: (a) starting with exophoric demonstratives and the shared physical space, and (b) abstracting away to their ongoing discourse representation with the emergence of anaphoric demonstratives, and finally (c) extending common ground to earlier experiences and

shared world knowledge with the emergence of definite articles (see Figure 3).

This particular instance of language change (from exophoric demonstratives, to anaphoric demonstratives and definite articles) has conceptual parallels in pragmatics and CSC. The use of demonstratives plays a key role in the automatization of attention monitoring and spatial perspective-taking during communication. Building on these foundational abilities, the use of anaphoric demonstratives and definite articles requires more sophisticated sociocognitive skills: monitoring ongoing discourse and earlier common ground requires, at a minimum, to be able to keep a record of what has been said and previously shared and, once fully developed, an understanding of what is known to the interlocutors in a conversation. Thus, the use of anaphoric demonstratives and definite articles ultimately trains speakers' *epistemic reasoning* (e.g., establishing whether your listener knows the name of your janitor, or you first need to introduce Fred in the conversation).

Methodological Challenges and Possible Approaches

Understanding the implications of language change for a speech community's CSC falls squarely within the scope of cultural evolutionary pragmatics. However, while the historical record offers documented evidence of language change, there is no parallel source of evidence for the evolution of CSC across generations of speakers. Given this methodological challenge, one way to study the effects of language change on CSC is the use of computational models and simulations (see, e.g., Hawkins et al., 2022; Kwisthout et al., 2008; Vogt & Coumans, 2003; Woensdregt et al., 2021).

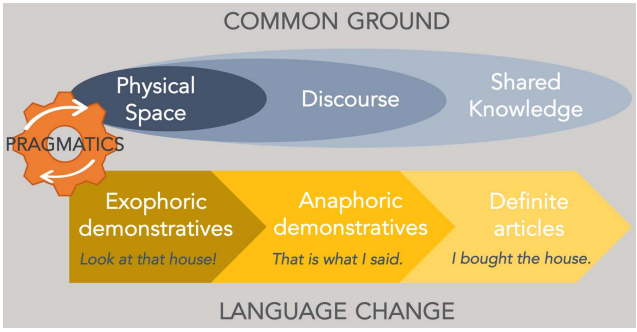
An alternative approach is to compare the use and processing of the different forms on a given grammaticalization pathway (e.g., the exophoric and anaphoric use of demonstratives; see Figure 3). Regarding the interpretation of "this" and "that" in written English, Çokal et al. (2014) argued that available theories were often based on an intuitive analogy between the spatial use of these forms in conversation and their anaphoric use in text. In this view, the proximal demonstrative would refer to more recently mentioned entities (analogous to objects that are closer to the speaker in space), while the distal form would refer to those antecedents that were mentioned earlier in the discourse (by analogy with the use of the distal demonstrative for far-away objects). However, the results of Çokal et al. (2014) ran counter to this intuitive analogy. Both demonstrative forms referred more frequently to recently mentioned antecedents than to earlier discourse referents, and their language production experiment revealed a small but statistically reliable difference contradicting the spatial analogy: the proximal demonstrative "this" was used more often to refer to early antecedents than the distal form "that." Çokal et al. (2014) interpreted these results in line with Cornish (2007), who argues that "this" is more deictically marked than "that," which would explain that the proximal demonstrative can bring about a *focal shift* toward earlier discourse segments.

Schumacher et al. (2015; Repp & Schumacher, 2022) used event-related brain potentials to better understand the processing of anaphoric demonstratives in German, by comparison with personal pronouns (for a related study using a text-continuation paradigm, see Fuchs & Schumacher, 2020; for a review of neurophysiological and neuroanatomical investigations of information structure, see Bornkessel-Schlesewsky & Schumacher, 2016). The ERPs revealed a biphasic N400—late positivity pattern at posterior electrodes for the demonstrative pronouns relative to the personal pronouns. Schumacher and colleagues interpret the N400 as an index of additional processing costs associated with more unexpected referential expression (with demonstrative pronouns being generally more unexpected than personal pronouns). The late positivity is taken to reflect the consequences of *attentional reorientation*: Since the demonstrative pronoun indicates a possible referential shift in the upcoming discourse (see Fossard et al., 2012; Fuchs & Schumacher, 2020), it induces updating of the discourse structure, which is normally indexed by a late positivity (for discussion and further empirical evidence, see Bornkessel-Schlesewsky & Schumacher, 2016; Burkhardt, 2006; Hirotani & Schumacher, 2011; Schumacher, 2009, 2012).

The results of Çokal et al. (2014) and Schumacher et al. (2015; Repp & Schumacher, 2022) therefore suggest that, contrary to the intuitive analogy between the spatial meaning of demonstratives and their anaphoric use in extended discourse, anaphoric demonstratives often initiate a *referential shift* toward a less prominent antecedent in the previous discourse (Fossard et al., 2012; Fuchs & Schumacher, 2020). This suggests that the anaphoric use of demonstratives might derive from the universal function of exophoric demonstratives to *redirect the listeners' visual attention* toward the intended referent (Rubio-Fernandez, 2021a, 2022; Woensdregt et al., 2022; see bottom row in Figure 2). Future studies should therefore directly compare the exophoric and anaphoric use of demonstratives in different languages to better understand their diachronic change and its implications for CSC—in particular, how their attentional reorientation function extends from physical space to the discourse model (see Figure 3).

A third methodological approach to investigate the effect of language change on the CSC of a speech community is to study these diachronic processes in real time. Psycholinguistics and language acquisition studies normally adopt a synchronic viewpoint, treating languages as fixed systems, but we know from diachronic linguistics that languages are constantly in flux (Heine & Kuteva, 2006). For instance, Givón (1981) characterizes the diachronic process whereby numerals are grammaticalized into

Figure 3
A Diachronic View of Common Ground



Note. Adapted from "Pragmatic Markers: The Missing Link Between Language and Theory of Mind," by P. Rubio-Fernandez, 2021a, *Synthese*, *199*(1), pp. 1125–1158 (https://doi.org/10.1007/s11229-020-02768-z). CC BY 4.0. See the online article for the color version of this figure.

indefinite markers (e.g., the change from the Old English form for "one" into the indefinite article "a" in modern English) as a universal pathway, after it was independently attested in Germanic, Romance, Mandarin, Sherpa, Hungarian, Neo-Aramaic, Persian, Turkish, various Amerindian, and Austronesian languages, in addition to being a hallmark of all Creole languages. The same process of grammaticalization has recently been identified in Chinese (Chen, 2004; Wong, 2016) and Polish (Hwaszcz & Kędzierska, 2018), and we recently investigated the case of Hindi, which, like Chinese and Polish, does not have an article system, but allows for the use of the numeral "ek" (one) to introduce new discourse referents (Dayal, 2004, 2018; Kachru, 1980, 2006; Sharma, 2005). These reference introduction uses correspond with the second stage of Heine's (1997) grammaticalization scale from numerals to indefinite articles.

Shukla et al. (2022) conducted a narrative-elicitation study comparing the acquisition of new/familiar markers in two languages without articles (Hindi and Mandinka) and two languages with articles (English and Spanish), and confirmed that the numeral "ek" is used to introduce new discourse characters in Hindi (Dayal, 2004; Sharma, 2005), although its use is not obligatory (cf. Dayal, 2018). In their first experiment, Hindi-speaking adults did not use "ek" for reference introduction as frequently as English- and Spanish-speaking adults used indefinite articles with the same function, which probably explains why 5-year-old Hindi speakers lagged behind English- and Spanish-speaking children of the same age in marking new referents. The second experiment further showed that by age 10 years, Hindi-speaking children use "ek" to introduce new characters at adult levels.

Similar patterns of results were observed with 5-year-olds, 10-year-olds and adults in Delhi and Gorakhpur, a city in the North Eastern state of Uttar Pradesh, suggesting that this linguistic phenomenon is more widespread than a mere dialectal variation around the Indian capital. Interestingly, however, adult speakers of Hindi differed in their use of the numeral "ek" for character introduction, with the highest rates being observed in Delhi, followed by Gorakhpur and finally Allahabad (a smaller city in Uttar Pradesh). These three groups varied in age and exposure to English, with younger adults with greater exposure to English using "ek" more frequently for reference introduction.

Assuming that an indefinite article is currently emerging in Hindi (Shukla et al., 2022), Chinese (Chen, 2004; Wong, 2016) and Polish (Hwaszcz & Kędzierska, 2018) from their numeral "one," future studies will have an invaluable opportunity to investigate the implications of this language change for the common-ground management of these speech communities. For example, will Hindi speakers increase their use of bare nouns as definite expressions signaling familiarity by contrast with the emerging use of "ek" to introduce new referents? (cf. similar contrasts in languages with only one article, such as Hebrew [definite] and Turkish [indefinite]). And if so, how would that affect Hindi speakers' discourse memory given what we know from the processing and recall of definite descriptions in English? (e.g., Clark & Sengul, 1979; Duff et al., 2011; Haviland & Clark, 1974; Irwin et al., 1982; Murphy, 1984). Addressing these and similar questions will be essential for our understanding of cultural evolutionary pragmatics (for a recent discussion of the importance of cross-linguistic and cross-cultural research for the advancement of cognitive science, see Blasi et al., 2022).

Closing Remarks

I will conclude with a recent quote from Leavens (2021, p. 10), in the context of the emergence of pointing in humans: "Human cultures are referential cultures." Here, I have further argued that in the referential nature of our cultures lies the deep connection between language and social cognition, which mutually reinforce each other in communication. In coming years, the study of cultural evolutionary pragmatics will hopefully shed light on how the acquisition, mature use and cultural evolution of reference systems sustain the coevolution and codevelopment of language and social cognition across cultures.

References

Agha, A. (2007). *Language and social relations*. Cambridge University Press.

Apperly, I. (2010). *Mindreaders: The cognitive basis of theory of mind*. Psychology Press. https://doi.org/10.4324/9780203833926

Bamberg, M. G. (1986). A functional approach to the acquisition of anaphoric relationships. *Linguistics*, *24*(1), 227–284. https://doi.org/10.1515/ling.1986.24.1.227

Bamberg, M. G. (1987). *The acquisition of narratives: Learning to use language*. Walter de Gruyter. https://doi.org/10.1515/9783110854190

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*(1), 1–22. https://doi.org/10.1006/jmla.1999.2667

Bard, E. G., & Aylett, M. P. (2005). Referential form, word duration, and modeling the listener in spoken dialogue. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 173–192). MIT Press.

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind* (pp. 233–251). Basil Blackwell.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Bradford/MIT Press. https://doi.org/10.7551/mitpress/4635.001.0001

Barth, D., Nicholas, E., Arka, I. W., Bergqvist, H., Forker, D., Gipper, S., Hodge, G., Kashima, E., Kasuga, Y., Kawakami, C., Kimoto, Y., Knuchel, D., Kogura, N., Kurabe, K., Mansfield, J., Narrog, H., Eka Pratiwi, D. P., van Putten, S., Senge, C., & Tykhostup, O. (2021). Language vs. individuals in cross-linguistic corpus typology. In G. Haig, S. Schnell, & F. Seifart (Eds.), *Doing corpus-based typology with spoken language data: State of the art* (pp. 1–56). Language Documentation & Conservation.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford University Press.

Beck, L., Kumschick, I. R., Eid, M., & Klann-Delius, G. (2012). Relationship between language competence and emotional competence in middle childhood. *Emotion*, *12*(3), 503–514. https://doi.org/10.1037/a0026320

Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, *1*(4), 557–560. https://doi.org/10.1017/S0140525X00076573

Benveniste, E. (1971). Subjectivity in language. In E. Meek (Ed.), *Problems in general linguistics* (pp. 223–230). University of Miami Press. (Original work published 1958).

Bergqvist, H., & Knuchel, D. (2019). Explorations of engagement: Introduction. *Open Linguistics*, *5*(1), 650–665. https://doi.org/10.1515/opli-2019-0036

Blakemore, D. (1987). *Semantic constraints on relevance*. Blackwell.

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences*, *26*(12), 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015

Boeg Thomsen, D., Theakston, A., Kandemirci, B., & Brandt, S. (2021). Do complement clauses really support false-belief reasoning? A longitudinal

study with English-speaking 2- to 3-year-olds. *Developmental Psychology*, *57*(8), 1210–1227. https://doi.org/10.1037/dev0001012

Bohnemeyer, J. (2018). Yucatec demonstratives in interaction: Spontaneous versus elicited data. In S. C. Levinson, S. Cutfield, M. J. Dunn, N. J. Enfield, & S. Meira (Eds.), *Demonstratives in cross-linguistic perspective* (pp. 176–205). Cambridge University Press. https://doi.org/10.1017/9781108333818.009

Bornkessel-Schlesewsky, I., & Schumacher, P. B. (2016). Towards a neurobiology of information structure. In C. Fry & S. Ishihara (Eds.), *The oxford handbook of information structure* (pp. 581–598). Oxford University Press.

Boroditsky, L. (2011). How language shapes thought. *Scientific American*, *304*(2), 62–65. https://doi.org/10.1038/scientificamerican0211-62

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, *2*(4), 331–350. https://doi.org/10.1016/0010-0285(71)90019-3

Brown-Schmidt, S., & Duff, M. C. (2016). Memory and common ground processes in language use. *Topics in Cognitive Science*, *8*(4), 722–736. https://doi.org/10.1111/tops.12224

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, *98*(2), 159–168. https://doi.org/10.1016/j.bandl.2006.04.005

Carmiol, A. M., Matthews, D., & Rodríguez-Villagra, O. A. (2018). How children learn to produce appropriate referring expressions in narratives: The role of clarification requests and modeling. *Journal of Child Language*, *45*(3), 736–752. https://doi.org/10.1017/S0305000917000381

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Blackwell Publishers. https://doi.org/10.1002/9780470754603

Chen, P. (2004). Identifiability and definiteness in Chinese. *Linguistics*, *42*(6), 1129–1184. https://doi.org/10.1515/ling.2004.42.6.1129

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Christiansen, M. H., & Chater, N. (2016). *Creating language: Integrating evolution, acquisition, and processing*. MIT Press. https://doi.org/10.7551/mitpress/10406.001.0001

Christiansen, M. H., & Kirby, S. E. (2003). *Language evolution*. Oxford University Press. https://doi.org/10.1093/acprof:Oso/9780199244843.001.0001

Clark, E. V. (1978). From gesture to word: On the natural history of deixis in language acquisition. In J. S. Bruner & A. Garton (Eds.), *Human growth and development* (pp. 85–120). Oxford University Press.

Clark, E. V. (2020). Conversational repair and the acquisition of language. *Discourse Processes*, *57*(5–6), 441–459. https://doi.org/10.1080/0163853X.2020.1719795

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62–81. https://doi.org/10.1016/j.jml.2003.08.004

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. Joshi, B. Webber, & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge University Press.

Clark, H. H., & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition*, *7*(1), 35–41. https://doi.org/10.3758/BF03196932

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7

Çokal, D., Sturt, P., & Ferreira, F. (2014). Deixis: This and that in written narrative discourse. *Discourse Processes*, *51*(3), 201–229. https://doi.org/10.1080/0163853X.2013.866484

Cole, P. M., Armstrong, L. M., & Pemberton, C. K. (2010). The role of language in the development of emotion regulation. In S. D. Calkins & M. A. Bell (Eds.), *Child development at the intersection of emotion and cognition* (pp. 59–77). American Psychological Association. https://doi.org/10.1037/12059-004

Comrie, B. (1989). *Language typology and language change*. Blackwell.

Cornish, F. (2007). English demonstratives: Discourse deixis and anaphora. A discourse-pragmatic account. In R. A. Nilsen, N. Aba Appiah Amfo, & K. Borthen (Eds.), *Interpreting utterances; pragmatics and its interfaces. Essays in honour of Thorstein Fretheim* (pp. 147–166). Novus Press.

Creissels, D. (2020). Grammaticalization in Manding languages. In W. Bisang & A. Malchukov (Eds.), *Grammaticalization scenarios: Areal patterns and cross-linguistic variation* (pp. 695–727). De Gruyter.

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, *70*(4), 853–865. https://doi.org/10.1111/1467-8624.00061

Dayal, V. (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy*, *27*(4), 393–450. https://doi.org/10.1023/B:LING.0000024420.80324.67

Dayal, V. (2018). (In)definiteness without articles: Diagnosis, analysis, implications. In G. Sharma & R. Bhatt (Eds.), *Trends in Hindi linguistics* (pp. 1–26). Mouton de Gruyter. https://doi.org/10.1515/9783110610796-001

De Cat, C. (2011). Information tracking and encoding in early L1: Linguistic competence vs. cognitive limitations. *Journal of Child Language*, *38*(4), 828–860. https://doi.org/10.1017/S030500091000036X

De Cat, C. (2013). Egocentric definiteness errors and perspective evaluation in preschool children. *Journal of Pragmatics*, *56*, 58–69. https://doi.org/10.1016/j.pragma.2012.08.002

de Villiers, J. (2007). The interface of language and theory of mind. *Lingua*, *117*(11), 1858–1878. https://doi.org/10.1016/j.lingua.2006.11.006

de Villiers, J. G. (1999). On acquiring the structural representations for false complements. In B. Hollebrandse (Ed.), *New perspectives on language acquisition* (pp. 125–136). UMOP.

de Villiers, P. A. (1974). Imagery and theme in recall of connected discourse. *Journal of Experimental Psychology*, *103*(2), 263–268. https://doi.org/10.1037/h0037608

Dediu, D., Cysouw, M., Levinson, S. C., Baronchelli, A., Christiansen, M. H., Croft, W., Evans, N., Garrod, S., Gray, R. D., Kandler, A., & Lieven, E. (2013). Cultural evolution of language. In P. J. Richerson & M. H. Christiansen (Eds.), *Cultural evolution: Society, technology, language, and religion* (Vol. 12, pp. 303–332). MIT Press. https://doi.org/10.7551/mitpress/9780262019750.003.0016

Dennett, D. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences*, *1*(4), 568–570. https://doi.org/10.1017/S0140525X00076664

Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics*, *17*(4), 463–489. https://doi.org/10.1515/COG.2006.015

Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, *25*(2), 108–127. https://doi.org/10.1016/j.newideapsych.2007.02.002

Diessel, H. (2013). Distance contrasts in demonstratives. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures*. Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/41

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. *PLOS ONE*, *10*(9), Article e0136100. https://doi.org/10.1371/journal.pone.0136100

Dor, D. (2015). *The instruction of imagination: Language as a social communication technology*. Oxford University Press. https://doi.org/10.1093/acprof:Oso/9780190256623.001.0001

Duff, M. C., Gupta, R., Hengst, J. A., Tranel, D., & Cohen, N. J. (2011). The use of definite references signals declarative memory: Evidence from patients with hippocampal amnesia. *Psychological Science*, *22*(5), 666–673. https://doi.org/10.1177/0956797611404897

Duranti, A. (2010). Husserl, intersubjectivity and anthropology. *Anthropological Theory*, *10*(1–2), 16–35. https://doi.org/10.1177/1463499610370517

Enfield, N. J. (2006). Social consequences of common ground. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality* (pp. 399–430). Routledge.

Enfield, N. J. (2015). Linguistic relativity from reference to agency. *Annual Review of Anthropology*, *44*(1), 207–224. https://doi.org/10.1146/annurev-anthro-102214-014053

Enfield, N. J., & Levinson, S. C. (2006). Introduction: Human sociality as a new interdisciplinary field. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality* (pp. 1–35). Routledge.

Evans, N. (2006). Who said polysynthetic languages avoid subordination? Multiple subordination strategies in Dalabon. *Australian Journal of Linguistics*, *26*(1), 31–58. https://doi.org/10.1080/07268600500531628

Evans, N. (2021). Social cognition in Dalabon. In D. Barth & N. Evans (Eds.), *Social Cognition Parallax Interview Corpus (SCOPIC)* (pp. 22–84). Language Documentation & Conservation.

Evans, N. (2022). *Words of wonder: Endangered languages and what they tell us*. Wiley-Blackwell. (Original work published 2009).

Evans, N., Bergqvist, H., & San Roque, L. (2018a). The grammar of engagement I: Framework and initial exemplification. *Language and Cognition*, *10*(1), 110–140. https://doi.org/10.1017/langcog.2017.21

Evans, N., Bergqvist, H., & San Roque, L. (2018b). The grammar of engagement II: Typology and diachrony. *Language and Cognition*, *10*(1), 141–170. https://doi.org/10.1017/langcog.2017.22

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(5), 429–448. https://doi.org/10.1017/S0140525X0999094X

Evers, S. (2020). *To "the" or not to "the": Cross-linguistic correlations between existing morphosyntax and the emergence of definite articles* [Unpublished doctoral dissertation]. The State University of New York at Buffalo.

Federici, A., Parma, V., Vicovaro, M., Radassao, L., Casartelli, L., & Ronconi, L. (2020). Anomalous perception of biological motion in autism: A conceptual review and meta-analysis. *Scientific Reports*, *10*(1), Article 4576. https://doi.org/10.1038/s41598-020-61252-3

Ferreira, V. S. (2019). A mechanistic framework for explaining audience design in language production. *Annual Review of Psychology*, *70*(1), 29–51. https://doi.org/10.1146/annurev-psych-122216-011653

Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*, *24*(1), 3–33. https://doi.org/10.3758/s13423-017-1236-5

Ford, C., & Fox, B. (1996). Interactional motivations for reference formulation: He had, this guy had, a beautiful, thirty-three olds. *Typological Studies in Language*, *33*, 145–168. https://doi.org/10.1075/tsl.33.06for

Forrester, M., & Cherington, S. (2009). The development of other-related conversational skills: A case study of conversational repair during the early years. *First Language*, *29*(2), 166–191. https://doi.org/10.1177/0142723708094452

Fossard, M., Garnham, A., & Cowles, H. W. (2012). Between anaphora and deixis … The resolution of the demonstrative noun phrase "that N". *Language and Cognitive Processes*, *27*(9), 1385–1404. https://doi.org/10.1080/01690965.2011.606668

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998. https://doi.org/10.1126/science.1218633

Fuchs, J., Domahs, U., & Kauschke, C. (2021). Information structure in language acquisition. Production and comprehension of (in)definite articles by German-speaking children. *Journal of Child Language*, *48*(1), 55–87. https://doi.org/10.1017/S0305000920000124

Fuchs, M., & Schumacher, P. B. (2020). Referential shift potential of demonstrative pronouns: Evidence from text continuation. In A. Næss, A. Margetts, & Y. Treis (Eds.), *Demonstratives in discourse* (pp. 185–213). Language Science Press.

Garde, M. (2008). Person reference, proper names and circumspection in Bininj Kun-wok conversation. In E. Mushin & B. Baker (Eds.), *Discourse and grammar in Australian languages* (pp. 203–232). John Benjamins. https://doi.org/10.1075/slcs.104.11gar

Garde, M. (2011). The forbidden gaze: The 1948 Wubarr ceremony performed for the American–Australian Scientific Expedition to Arnhem Land. In M. Thomas & M. Neale (Eds.), *Exploring the legacy of the 1948 Arnhem Land expedition* (pp. 403–421). ANU Press. https://doi.org/10.22459/ELALE.06.2011.20

Garrod, S. C., & Sanford, A. J. (1982). The mental representation of discourse in a focussed memory system: Implications for the interpretation of anaphoric noun phrases. *Journal of Semantics*, *1*(1), 21–41. https://doi.org/10.1093/jos/1.1.21

Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind: Advances in the study of language and thought*. MIT Press. https://doi.org/10.7551/mitpress/4117.001.0001

Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing. *Discourse Processes*, *39*(2), 225–242. https://doi.org/10.1080/0163853X.2005.9651681

Givón, T. (1981). On the development of the numeral "one" as an indefinite marker. *Folia Linguistica Historica*, *15*(1), 35–54. https://doi.org/10.1515/flih.1981.2.1.35

Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, *30*(1), 5–56. https://doi.org/10.1515/ling.1992.30.1.5

Givón, T. (2001). *Syntax: An introduction* (Vol. 1). John Benjamins Publishing.

Glanzberg, M. (2014). Explanation and partiality in semantic theory. In A. Burgess & B. Sherman (Eds.), *Metasemantics: New essays on the foundations of meaning* (pp. 259–292). Oxford University Press. https://doi.org/10.1093/acprof:Oso/9780199669592.003.0011

Goody, E. (1995). Introduction: Some implications of a social origin of intelligence. In E. Goody (Ed.), *Social intelligence and interaction* (pp. 1–33). Cambridge University Press. https://doi.org/10.1017/CBO9780511621710.003

Gori, M., Schiatti, L., & Amadeo, M. B. (2021). Masking emotions: Face masks impair how we read emotions. *Frontiers in Psychology*, *12*, Article 669432. https://doi.org/10.3389/fpsyg.2021.669432

Greenberg, J. H. (1978). How does a language acquire gender markers. *Universals of Human Language*, *3*, 47–82.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and semantics: Vol. 3. *speech acts* (pp. 41–58). Academic Press.

Grimshaw, J. (2005). *Words and structure*. CSLI Publications.

Gundel, J. K., & Johnson, K. (2013). Children's use of referring expressions in spontaneous discourse: Implications for theory of mind development. *Journal of Pragmatics*, *56*, 43–57. https://doi.org/10.1016/j.pragma.2013.04.003

Gundel, J. K., Ntelitheos, D., & Kowalsky, M. (2007). Children's use of referring expressions: Some implications for theory of mind. *ZAS Papers in Linguistics*, *48*, 1–21. https://doi.org/10.21248/zaspil.48.2007.351

Hacquard, V., & Lidz, J. (2022). On the acquisition of attitude verbs. *Annual Review of Linguistics*, *8*(1), 193–212. https://doi.org/10.1146/annurev-linguistics-032521-053009

Hanks, W. (2005). Explorations in the deictic field. *Current Anthropology*, *46*(2), 191–220. https://doi.org/10.1086/427120

Hanks, W. F. (2009). Fieldwork on deixis. *Journal of Pragmatics*, *41*(1), 10–24. https://doi.org/10.1016/j.pragma.2008.09.003

Harman, G. (1978). Studying the chimpanzees' theory of mind. *Behavioral and Brain Sciences*, *1*(4), 576–577. https://doi.org/10.1017/S0140525X00076743

Harris, D. W. (2022). Semantics without semantic content. *Mind & Language*, *37*(3), 304–328. https://doi.org/10.1111/mila.12290

Haviland, J. B. (2007). Person reference in Tzotzil gossip: Referring dupliciter. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 226–252).

Cambridge University Press. https://doi.org/10.1017/CBO9780511486746.011

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 512–521. https://doi.org/10.1016/S0022-5371(74)80003-4

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

Heasman, B., & Gillespie, A. (2019). Neurodivergent intersubjectivity: Distinctive features of how autistic people create shared understanding. *Autism: An International Journal of Research and Practise*, *23*(4), 910–921. https://doi.org/10.1177/1362361318785172

Heine, B. (1997). *Cognitive foundations of grammar*. Oxford University Press.

Heine, B., & Kuteva, T. (2006). *The changing languages of Europe*. Oxford University Press on Demand.

Heiphetz, L., Spelke, E. S., Harris, P. L., & Banaji, M. R. (2013). The development of reasoning about beliefs: Fact, preference, and ideology. *Journal of Experimental Social Psychology*, *49*(3), 559–565. https://doi.org/10.1016/j.jesp.2012.09.005

Heritage, J. (2007). Intersubjectivity and progressivity in person (and place) reference. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 255–280). Cambridge University Press. https://doi.org/10.1017/CBO9780511486746.012

Heyes, C. (2018). *Cognitive gadgets*. Harvard University Press.

Heyes, C. (2019a). Précis of *Cognitive gadgets*: The cultural evolution of thinking. *Behavioral and Brain Sciences*, *42*(e169), 1–58. https://doi.org/10.1017/S0140525X18002145

Heyes, C. (2019b). Testing cognitive gadgets. *Mind & Language*, *34*(4), 551–559. https://doi.org/10.1111/mila.12253

Heyes, C. (2020). Psychological mechanisms forged by cultural evolution. *Current Directions in Psychological Science*, *29*(4), 399–404. https://doi.org/10.1177/0963721420917736

Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*(6190), Article 1243091. https://doi.org/10.1126/science.1243091

Hickmann, M., & Hendriks, H. (1999). Cohesion and anaphora in children's narratives: A comparison of English, French German, and Mandarin Chinese. *Journal of Child Language*, *26*(2), 419–452. https://doi.org/10.1017/S0305000999003785

Hickmann, M., Schimke, S., & Colonna, S. (2015). From early to late mastery of reference: Multifunctionality and linguistic diversity. In L. Serratrice & S. E. M. Allen (Eds.), *The acquisition of reference, TiLAR 15* (pp. 181–211). John Benjamins Publishing Company.

Hirotani, M., & Schumacher, P. B. (2011). Context and topic marking affect distinct processes during discourse comprehension in Japanese. *Journal of Neurolinguistics*, *24*(3), 276–292. https://doi.org/10.1016/j.jneuroling.2010.09.007

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Hupet, M., & Chantraine, Y. (1992). Changes in repeated references: Collaboration or repetition effects? *Journal of Psycholinguistic Research*, *21*(6), 485–496. https://doi.org/10.1007/BF01067526

Hwaszcz, K., & Kędzierska, H. (2018). The rise of an indefinite article in Polish: An appraisal of its grammaticalisation stage. *Studies in Polish Linguistics*, *13*, 145–166. https://doi.org/10.4467/23005920SPL.18.005.8744

Irwin, D. E., Bock, J. K., & Stanovich, K. E. (1982). Effects of information structure cues on visual word processing. *Journal of Verbal Learning and Verbal Behavior*, *21*(3), 307–325. https://doi.org/10.1016/S0022-5371(82)90637-5

Jackendoff, R. (1983). *Semantics and cognition*. MIT Press.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, and evolution*. Oxford University Press. https://doi.org/10.1093/acprof:Oso/9780198270126.001.0001

Jacoby, N., & Fedorenko, E. (2020). Discourse-level comprehension engages medial frontal theory of mind brain regions even for expository texts. *Language, Cognition and Neuroscience*, *35*(6), 780–796. https://doi.org/10.1080/23273798.2018.1525494

Jakobson, R. (2013). On linguistic aspects of translation. In R. A. Brown (Ed.), *On translation* (pp. 232–239). Harvard University Press. (Original work published 1959).

Jara-Ettinger, J., & Rubio-Fernandez, P. (2021). Quantitative mental state attributions in language understanding. *Science Advances*, *7*(47), Article eabj0970. https://doi.org/10.1126/sciadv.abj0970

Jara-Ettinger, J., & Rubio-Fernandez, P. (2022). The social basis of referential communication: Speakers construct physical reference based on listeners' expected visual search. *Psychological Review*, *129*(6), 1394–1413. https://doi.org/10.1037/rev0000345.

Jones, D. (2010). Human kinship, from conceptual structure to grammar. *Behavioral and Brain Sciences*, *33*(5), 367–404. https://doi.org/10.1017/S0140525X10000890

Kachru, Y. (1980). *Aspects of Hindi grammar*. Manohar Publications.

Kachru, Y. (2006). *Hindi*. John Benjamins Publishing. https://doi.org/10.1075/loall.12

Keesing, R. M. (1981). *Cultural anthropology: A contemporary perspective*. Holt, Rinehard & Winston.

Kita, S. (Ed.). (2003). *Pointing: Where language, culture, and cognition meet*. Psychology Press. https://doi.org/10.4324/9781410607744

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843. https://doi.org/10.1038/nrn1533

Kwisthout, J., Vogt, P., Haselager, P., & Dijkstra, T. (2008). Joint attention and language evolution. *Connection Science*, *20*(2–3), 155–171. https://doi.org/10.1080/09540090802091958

Leavens, D. A. (2021). The referential problem space revisited: An ecological hypothesis of the evolutionary and developmental origins of pointing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *12*(4), Article e1554. https://doi.org/10.1002/wcs.1554

Lesgold, A. M. (1972). Pronominalization: A device for unifying sentences in memory. *Journal of Verbal Learning and Verbal Behavior*, *11*(3), 316–323. https://doi.org/10.1016/S0022-5371(72)80093-8

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Levinson, S. C. (2006). Cognition at the heart of human interaction. *Discourse Studies*, *8*(1), 85–93. https://doi.org/10.1177/1461445606059557

Levinson, S. C. (2018). Demonstratives: Patterns in diversity. In S. C. Levinson, S. Cutfield, M. Dunn, N. Enfield, & S. Meira (Eds.), *Demonstratives in cross-Linguistic perspective* (pp. 1–42). Cambridge University Press. https://doi.org/10.1017/9781108333818.002

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, *12*(2), 264–271. https://doi.org/10.1111/j.1467-7687.2008.00758.x

Liebal, K., Carpenter, M., & Tomasello, M. (2011). Young children's understanding of markedness in non-verbal communication. *Journal of Child Language*, *38*(4), 888–903. https://doi.org/10.1017/S0305000910000383

Liszkowski, U. (2018). Emergence of shared reference and shared minds in infancy. *Current Opinion in Psychology*, *23*, 26–29. https://doi.org/10.1016/j.copsyc.2017.11.003

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development*, *74*(4), 1130–1144. https://doi.org/10.1111/1467-8624.00597

Long, M., Moore, I., Mollica, F., & Rubio-Fernandez, P. (2021). Contrast perception as a visual heuristic in the formulation of referential expressions. *Cognition*, *217*, Article 104879. https://doi.org/10.1016/j.cognition.2021.104879

Lupyan, G., Abdel Rahman, R., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences*, *24*(11), 930–944. https://doi.org/10.1016/j.tics.2020.08.005

Lyons, C. (1999). *Definiteness*. Cambridge University Press. https://doi.org/10.1017/CBO9780511605789

Majid, A. (2021). Human olfaction at the intersection of language, culture, and biology. *Trends in Cognitive Sciences*, *25*(2), 111–123. https://doi.org/10.1016/j.tics.2020.11.005

Malle, B. F. (2002). The relation between language and Theory of Mind in development and evolution. In T. Givón & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). Benjamins. https://doi.org/10.1075/tsl.53.14mal

Matsui, T., Yamamoto, T., & McCagg, P. (2006). On the role of language in children's early understanding of others as epistemic beings. *Cognitive Development*, *21*(2), 158–173. https://doi.org/10.1016/j.cogdev.2005.10.001

Matthews, D., Lieven, E., & Tomasello, M. (2007). How toddlers and preschoolers learn to uniquely identify referents for others: A training study. *Child Development*, *78*(6), 1744–1759. https://doi.org/10.1111/j.1467-8624.2007.01098.x

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*(2), 622–646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Mithun, M. (1984). How to avoid subordination. *Berkeley Linguistic Society*, *10*, 493–509.

Moll, H., Carpenter, M., & Tomasello, M. (2007). Fourteen-month-olds know what others experience only in joint engagement. *Developmental Science*, *10*(6), 826–835. https://doi.org/10.1111/j.1467-7687.2007.00615.x

Moll, H., & Kadipasaoglu, D. (2013). The primacy of social over visual perspective-taking. *Frontiers in Human Neuroscience*, *7*, Article 558. https://doi.org/10.3389/fnhum.2013.00558

Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what "we" have shared in a special way. *Infancy*, *13*(1), 90–101. https://doi.org/10.1080/15250000701779402

Moll, H., & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, *43*(2), 309–317. https://doi.org/10.1037/0012-1649.43.2.309

Moore, R. (2021). The cultural evolution of mind-modelling. *Synthese*, *199*(1), 1751–1776. https://doi.org/10.1007/s11229-020-02853-3

Murphy, G. L. (1984). Establishing and accessing referents in discourse. *Memory & Cognition*, *12*(5), 489–497. https://doi.org/10.3758/BF03198311

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, *67*(2), 659–677. https://doi.org/10.2307/1131839

Ornaghi, V., & Grazzani, I. (2013). The relationship between emotional-state language and emotion understanding: A study with school-age children. *Cognition and Emotion*, *27*(2), 356–366. https://doi.org/10.1080/02699931.2012.711745

Özyürek, A. (1998). An analysis of the basic meaning of Turkish demonstratives in face-to-face conversational interaction. In S. Santi, I. Guaitella, C. Cave, & G. Konopczynski (Eds.), *Oralite et Gestualite: Communication multimodale, interaction* (pp. 609–614). L'Harmattan.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *369*(1651), Article 20130296. https://doi.org/10.1098/rstb.2013.0296

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*(1), 155–184. https://doi.org/10.1016/j.cognition.2008.02.007

Paunov, A. M., Blank, I. A., & Fedorenko, E. (2019). Functionally distinct language and theory of mind networks are synchronized at rest and during language comprehension. *Journal of Neurophysiology*, *121*(4), 1244–1265. https://doi.org/10.1152/jn.00619.2018

Paunov, A. M., Blank, I. A., Jouravlev, O., Mineroff, Z., Gallée, J., & Fedorenko, E. (2022). Differential tracking of linguistic vs. mental state content in naturalistic stimuli by language and theory of mind (ToM) brain networks. *Neurobiology of Language*, *3*(3), 413–440. https://doi.org/10.1162/nol_a_00071

Pérez, D. I., & Gomila, A. (2021). *Social cognition and the second person in human interaction*. Routledge. https://doi.org/10.4324/9781003133155

Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. A. Frensch & R. Schwarzer (Eds.), *Cognition and neuropsychology* (pp. 255–276). Psychology Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. MIT Press.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. Viking.

Planer, R., & Sterelny, K. (2021). *From signal to symbol: The evolution of language*. MIT Press. https://doi.org/10.7551/mitpress/13906.001.0001

Pleyer, M., & Winters, J. (2014). Integrating cognitive linguistics and language evolution research. *Theoria et Historia Scientiarum*, *11*, 19–44. https://doi.org/10.12775/ths-2014-002

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*(2), 127–152. https://doi.org/10.1080/17405620344000022

Power, R. J., & Dal Martello, M. F. (1986). The use of the definite and indefinite articles by Italian preschool children. *Journal of Child Language*, *13*(1), 145–154. https://doi.org/10.1017/S0305000900000350

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Prunty, J. E., Jenkins, R., Qarooni, R., & Bindemann, M. (2022). Ingroup and outgroup differences in face detection. *British Journal of Psychology*. Advance online publication. https://doi.org/10.1111/bjop.12588

Pyke, A., West, R. L., & LeFevre, J. A. (2007a). *How readers retrieve referents for nouns in real time: A memory-based model of context effects on referent accessibility* [Conference session]. Proceedings of the 8th International Conference on Cognitive Modeling, New York, United States.

Pyke, A., West, R. L., & LeFevre, J. A. (2007b). On-line reference assignment for anaphoric and non-anaphoric nouns: A unified, memory-based model in ACT-R. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 1403–1408). Cognitive Science Society.

Rakoczy, H. (2017). In defense of a developmental dogma: Children acquire propositional attitude folk psychology around age 4. *Synthese*, *194*(3), 689–707. https://doi.org/10.1007/s11229-015-0860-8

Repp, M., & Schumacher, P. (2022). *Using naturalistic stimuli to investigate online pronoun processing in German* [Paper presentation]. XPrag 2022 Conference, Pavia, Italy.

Richardson, H., Koster-Hale, J., Caselli, N., Magid, R., Benedict, R., Olson, H., Pyers, J., & Saxe, R. (2020). Reduced neural selectivity for mental states in deaf children with delayed exposure to sign language. *Nature Communications*, *11*(1), Article 3246. https://doi.org/10.1038/s41467-020-17004-y

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nature Communications*, *9*(1), Article 1027. https://doi.org/10.1038/s41467-018-03399-2

Robbins, J., & Rumsey, A. (2008). Introduction: Cultural and linguistic anthropology and the opacity of other minds. *Anthropological Quarterly*, *81*(2), 407–420. https://doi.org/10.1353/anq.0.0005

Rubio-Fernandez, P. (2019). Overinformative speakers are cooperative: Revisiting the Gricean Maxim of Quantity. *Cognitive Science*, *43*(11), Article e12797. https://doi.org/10.1111/cogs.12797

Rubio-Fernandez, P. (2021a). Pragmatic markers: The missing link between language and theory of mind. *Synthese*, *199*(1), 1125–1158. https://doi.org/10.1007/s11229-020-02768-z

Rubio-Fernandez, P. (2021b). Color discriminability makes over-specification efficient: Theoretical analysis and empirical evidence. *Humanities & Social Sciences Communications*, *8*(1), 1–15. https://doi.org/10.1057/s41599-021-00818-6

Rubio-Fernandez, P. (2022). Demonstrative systems: From linguistic typology to social cognition. *Cognitive Psychology*, *139*, Article 101519. https://doi.org/10.1016/j.cogpsych.2022.101519

Rubio-Fernandez, P., & Jara-Ettinger, J. (2021). *A cross-cultural study of the use and comprehension of color words: English vs. Mandinka* [Paper presentation]. Conference for Architectures and Mechanisms for Language Processing (AMLaP 2021), Paris, France.

Rubio-Fernandez, P., Mollica, F., & Jara-Ettinger, J. (2021). Speakers and listeners exploit word order for communicative efficiency: A cross-linguistic investigation. *Journal of Experimental Psychology: General*, *150*(3), 583–594. https://doi.org/10.1037/xge0000963

Rubio-Fernandez, P., Shukla, V., Bhatia, V., Ben-Ami, S., & Sinha, P. (2022). Head turning is an effective cue for gaze following: Evidence from newly sighted individuals, school children and adults. *Neuropsychologia*, *174*, Article 108330. https://doi.org/10.1016/j.neuropsychologia.2022.108330

Rubio-Fernández, P., Mollica, F., Oraa Ali, M., & Gibson, E. (2019). How do you know that? Automatic belief inferences in passing conversation. *Cognition*, *193*, Article 104011. https://doi.org/10.1016/j.cognition.2019.104011

Sacks, H., & Schegloff, E. A. (2007). Two preferences in the organization ofreference to persons in conversation and their interaction. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 23–28). Cambridge University Press. (Original work published 1973).

Saylor, M. M., & Ganea, P. (2007). Infants interpret ambiguous requests for absent objects. *Developmental Psychology*, *43*(3), 696–704. https://doi.org/10.1037/0012-1649.43.3.696

Schneider, R. M., Yurovsky, D., & Frank, M. (2015). *Large-scale investigations of variability in children's first words* [Conference session]. Proceedings of the 2015 Meeting of the Cognitive Science Society, Pasadena, California, United States.

Schumacher, P. B. (2009). Definiteness marking shows late effects during discourse processing: Evidence from ERPs. In S. Lalitha Devi, A. Branco, & R. Mitkov (Eds.), *Anaphora processing and applications* (pp. 91–106). Springer. https://doi.org/10.1007/978-3-642-04975-0_8

Schumacher, P. B. (2012). Context in neurolinguistics: Time-course data from electrophysiology. In R. Finkbeiner, J. Meibauer, & P. B. Schumacher (Eds.), *What is a context? Linguistic approaches and challenges* (pp. 33–54). John Benjamins. https://doi.org/10.1075/la.196.05sch

Schumacher, P. B., Backhaus, J., & Dangl, M. (2015). Backward- and forward-looking potential of anaphors. *Frontiers in Psychology*, *6*, Article 1746. https://doi.org/10.3389/fpsyg.2015.01746

Scott-Phillips, T. (2014). *Speaking our Minds*. Palgrave Macmillan.

Sharma, D. (2005). Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition*, *27*(4), 535–566. https://doi.org/10.1017/S0272263105050242

Shukla, V., Long, M., & Rubio-Fernandez, P. (2022). Children's acquisition of new/given markers in English, Hindi, Mandinka and Spanish: Exploring the effect of optionality during grammaticalization. *Glossa Psycholinguistics*, *1*(1), Article 13. https://doi.org/10.5070/G6011120

Skilton, A. (2022). Learning speaker-and addressee-centered demonstratives in Ticuna. *Journal of Child Language*. Advance online publication. https://doi.org/10.1017/S0305000922000101

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking". In J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge University Press.

Smith, K., & Kirby, S. (2008). Cultural evolution: Implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *363*(1509), 3591–3603. https://doi.org/10.1098/rstb.2008.0145

Smith, L. B. (2013). It's all connected: Pathways in visual object recognition and early noun learning. *American Psychologist*, *68*(8), 618–629. https://doi.org/10.1037/a0034185

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*(4), 325–336. https://doi.org/10.1016/j.tics.2018.02.004

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Basil Blackwell. (Original work published 1986).

Stern, D. (2005). Intersubjectivity. In E. S. Person, A. M. Cooper, & G. O. Gabbard (Eds.), *The American psychiatric publishing textbook of psychoanalysis* (pp. 77–92). American Psychiatric Publishing.

Stivers, T., Enfield, N. J., & Levinson, S. C. (2007). Person reference in interaction. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 1–20). Cambridge University Press.

Streubel, B., Gunzenhauser, C., Grosse, G., & Saalbach, H. (2020). Emotion-specific vocabulary and its contribution to emotion understanding in 4- to 9-year-old children. *Journal of Experimental Child Psychology*, *193*, Article 104790. https://doi.org/10.1016/j.jecp.2019.104790

Talmy, L. (2000). The relation of grammar to cognition. In L. Talmy (Ed.), *Toward a cognitive semantics* (Vol. I, pp. 21–96). Cambridge University Press. https://doi.org/10.7551/mitpress/6847.003.0003

Tamariz, M., & Kirby, S. (2016). The cultural evolution of language. *Current Opinion in Psychology*, *8*, 37–43. https://doi.org/10.1016/j.copsyc.2015.09.003

Taumoepeau, M., & Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child Development*, *77*(2), 465–481. https://doi.org/10.1111/j.1467-8624.2006.00882.x

Taumoepeau, M., & Ruffman, T. (2008). Stepping stones to others' minds: Maternal talk relates to child mental state language and emotion understanding at 15, 24, and 33 months. *Child Development*, *79*(2), 284–302. https://doi.org/10.1111/j.1467-8624.2007.01126.x

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT press.

Thelen, E., & Smith, L. B. (2007). Dynamic systems theories. In R. M. Lerner (Ed.), Handbook of child psychology: Vol. 1. *Theoretical models of human development* (6th ed., pp. 258–312). Wiley.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Tomasello, M. (2008). *Origins of human communication*. MIT Press. https://doi.org/10.7551/mitpress/7551.001.0001

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691. https://doi.org/10.1017/S0140525X05000129

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology*, *39*(5), 906–912. https://doi.org/10.1037/0012-1649.39.5.906

Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 321–348). Cambridge University Press.

Ünal, E., & Papafragou, A. (2018). The relation between language and mental state reasoning. In J. Proust & M. Frontier (Eds.), *Metacognitive diversity: An interdisciplinary approach* (pp. 153–169). Oxford University Press.

Ünal, E., & Papafragou, A. (2020). Relations between language and cognition: Evidentiality and sources of knowledge. *Topics in Cognitive Science*, *12*(1), 115–135. https://doi.org/10.1111/tops.12355

Verhagen, A. (2007). *Constructions of intersubjectivity: Discourse, syntax, and cognition*. Oxford University Press. https://doi.org/10.1093/acprof:Oso/9780199226702.001.0001

Verhagen, A. (2015). Grammar and cooperative communication. In E. Dąbrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 232–252). De Gruyter Mouton. https://doi.org/10.1515/9783110292022-012

Vogt, P., & Coumans, H. (2003). Investigating social interaction strategies for bootstrapping lexicon development. *Journal of Artificial Societies and Social Simulation*, *6*(1), 1–4. https://www.jasss.org/6/1/4.html

Warden, D. (1976). The influence of context on children's use of identifying expressions and references. *British Journal of Psychology*, *67*(1), 101–112. https://doi.org/10.1111/j.2044-8295.1976.tb01501.x

Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press. https://doi.org/10.1093/acprof:Oso/9780199334919.001.0001

Whiten, A. (2000). Chimpanzee cognition and the question of mental re-representation. In D. Sperber (Ed.), *Metarepresentation: A multidisciplinary perspective* (pp. 139–167). Oxford University Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Woensdregt, M., Cummins, C., & Smith, K. (2021). A computational model of the cultural co-evolution of language and mindreading. *Synthese*, *199*(1), 1347–1385. https://doi.org/10.1007/s11229-020-02798-7

Woensdregt, M., Jara-Ettinger, J., & Rubio-Fernandez, P. (2022). Language universals rely on social cognition: Computational models of the use of this and that to redirect the receiver's attention. In *Proceedings of the 44th annual meeting of the cognitive science society (CogSci 2022)* (pp. 1382–1388). Cognitive Science Society.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253–265. https://doi.org/10.1002/wcs.104

Wong, A. L. (2016). Indefinite markers, grammaticalization, and language contact phenomena in Chinese. *Proceedings of the Linguistic Society of America*, *1*, 9–1. https://doi.org/10.3765/plsa.v1i0.3702

Wong, A. M. Y., & Johnston, J. R. (2004). The development of discourse referencing in Cantonese-speaking children. *Journal of Child Language*, *31*(3), 633–660. https://doi.org/10.1017/S030500090400604X

Zahavi, D., & Overgaard, S. (2013). Intersubjectivity. In H. LaFollette (Ed.), *The international encyclopedia of ethics* (pp. 2755–2765). Wiley-Blackwell. https://doi.org/10.1002/9781444367072.wbiee274

Received September 6, 2022
Revision received January 3, 2023
Accepted January 3, 2023 ■