CellPress
OPEN ACCESS

Trends in
Cognitive Sciences

Review

Spatial communication systems and action

Kenny R. Coventry [1,*] and Holger Diessel [2]

Spatial cognition is fundamental to our species. One might therefore expect that spatial communication systems would have evolved to make common distinctions. However, many have argued that spatial communication systems exhibit considerable cross-linguistic diversity, challenging the view that space structures language. We review recent work on spatial communication that merits revisiting the relationship between language and space. We provide a framework that places action as the driver of spatial communication systems across languages, in which spatial demonstratives – the earliest spatial terms – play a fundamental role in honing attention and theory of mind capacities that are crucial for language and cognition more broadly. We discuss how demonstratives emerged early in language evolution to serve a combination of spatial, social, and functional needs.

Highlights

Recent advances indicate that action is an important constraint on spatial communication systems and may provide a universal structuring tool for communicating about space across languages.

Spatial demonstratives – the earliest spatial terms – embody action and attentional components that may play a fundamental role in the development of social cognition and theory of mind more broadly.

The occurrence of demonstratives early in language evolution represents a confluence of gesture, social and spatial constraints unifying theories of language evolution.

Spatial communication and linguistic diversity

Navigating and locating objects in the environment is essential for all species. In contrast to other animals, however, humans have a unique capacity to communicate flexibly about the locations of people, objects, and places [1]. Expressions such as 'that spoon', 'to the left of the cup', 'in front of the cinema', and 'up and away' direct the attention of a hearer to a location, places, or events that can be in the immediate vicinity (e.g., on the table in front of speaker and hearer) or remote from both speaker and hearer. It has long been noted that spatial terms also crop up in non-spatial contexts, including temporal expressions ('see you in 5 minutes', 'the basketball game was over', 'that event'), and expressions of emotion ("I'm on top of the world", "I'm feeling under the weather"), leading some to argue that space is pivotal to understanding and organizing non-spatial domains [2–4]. More broadly, many have championed space as the pivotal structuring tool for both linguistic [2,5–9] and non-linguistic cognition [10,11].

The prevalence of space as a structuring tool for human cognition often comes with the assumption that spatial concepts are universal. Consistent with the Fodorian view that languages encode the categories one thinks in [12], one might assume that learning a language is simply a matter of finding out how the local language expresses the universal spatial concepts that one already has [13,14]. Primary among proposed universal concepts is the idea, originally expressed by Kant [15], that space is conceived of in relation to the body of the perceiver. This 'premise of egocentric primacy' ([16], p.125) has dominated theories of cognitive development and language acquisition (e.g., [17]) and theories of spatial cognition ever since [18,19].

However, the twin assumptions of the primacy of the egocentric perspective, on the one hand, and the universality of spatial language and spatial concepts, on the other, have been challenged by a resurgence of earlier views of **linguistic relativity** (see Glossary) associated with Sapir and Whorf [20,21]. The main reason for the rejection of universal semantic categories (or indeed, any linguistic universals at all [22]) is the recognition of considerable cross-linguistic variation in spatial communication systems. There is much evidence from **spatial adpositional** categories (i.e., prepositions and postpositions [23,24]) that languages carve up space in different ways,

[1]University of East Anglia, Norwich NR4 7TJ, UK
[2]Friedrich Schiller University Jena, 07743 Jena, Germany

*Correspondence:
k.coventry@uea.ac.uk (K.R. Coventry).

and in ways that directly challenge the premise of egocentric primacy. Speakers of a range of languages (e.g., Tzeltal, Guugu Yimitrr, Haillom) employ allocentric/absolute/geocentric **reference frames** even when describing the location of objects in tabletop space [25–27]. When describing the relative locations of objects on a table, for example, English speakers employ expressions such as 'the pencil is to the left of the eraser'; by contrast, Tzeltal speakers employ cardinal directions (e.g., 'the pencil is north of the eraser'). Analyses of such languages were pivotal in the development of **Levinson's typology of spatial reference frames** as expressed in language [26,28] that dispenses with the **egocentric/allocentric distinction**, and in so doing diminishes the role of the ego in spatial description. Moreover, a range of studies point to a relationship between the dominant spatial reference system expressed in individual languages and the nonlinguistic spatial cognition of speakers of those languages [25,29–31], supporting the view that language structures space as much as space structures language (*cf* [32,33]). These dual views – the rejection of the egocentric frame as a default in spatial language and spatial cognition, and the associated claim that semantic universals do not exist for spatial communication – have now become the new orthodoxy for many in cognitive science.

In this review article we reconsider the relationship between spatial language and spatial cognition through the lens of the most important spatial terms in language – spatial demonstratives. These terms have been largely neglected in debates regarding diversity in spatial communication and spatial cognition, and the literature has almost exclusively focused on spatial adpositions (and their equivalents) across languages. We show that recent research advances in understanding demonstratives from a cross-linguistic perspective challenge the view that there are no common constraints on spatial communication systems across languages, and that the premise of egocentric primacy should be rejected. We show that spatial demonstratives involve intimate connections between action, space, and attention, thus uniting cognitive and social processes that make human communication special. In placing action at the core of spatial communication, we reinstate the importance of semantic universals of space that situate the ego at the center of the action.

Spatial demonstrative systems across languages

Spatial demonstratives are a class of expressions comprising nominal (e.g., 'this' and 'that' in English) and adverbial forms (e.g., 'here' and 'there' in English) that are closely related. Although there is continued debate regarding their precise function in communication, it is generally recognized that they play a key role in both coordinating joint attention and directing attention to an object or objects in space. The former function is closely related to pointing behavior – one of the first means of intentional communication [34–36]. The directing function of demonstratives is associated with **deixis** [37,38] – in common with some other spatial terms (e.g., the so-called 'projective' spatial adpositions such as 'to the left/right of'), demonstrative meaning is determined by the context of use and by the spatial perspective one takes.

For many reasons, it can be argued that demonstratives should be the starting point when considering spatial communication and communication more broadly. First, although not all languages have spatial adpositions [26], typological studies sampling hundreds of languages [39–41] have not found a single language that lacks demonstratives [42], making them a likely universal linguistic category and countering views favoring extreme diversity [22]. Second, there is evidence that demonstratives are among the earliest forms in language evolution and may have played a central role in the development of language (Box 1) [43]. Third, there is also evidence that demonstratives are among the earliest words to appear in language acquisition across languages [44,45] (Box 2). Last, spatial demonstratives can be set apart from other forms of verbal spatial communication because they are the spatial terms most closely linked to gesture and eye

Glossary

Deixis: Bühler originally defined deixis as applied to linguistic terms that are semantically contingent on a particular point of reference, but the term has also been used more broadly to refer to terms that are fully dependent on context to be understood.
Dorsal 'where' system: this is the visual system that extends from primary visual cortex in the occipital lobe to the posterior parietal cortex, also known as the 'where' or 'how' stream owing to its role in locating objects in space and guiding action to objects (as distinct from the ventral 'what' system).
Egocentric/allocentric distinction: this distinction is between a viewer-centred perspective (i.e., 'the car is to the right of the truck from my point of view'; 'the dog is in front of me', etc.) and a perspective based on external coordinates that can be centered on another object, another person, or on fixed points in the environment.
Levinson's typology of reference frames: Levinson proposed a three-way distinction between intrinsic, relative, and absolute frames. The intrinsic frame is a binary relationship between a located object and a reference object from which a viewpoint emanates (e.g., 'the cup is to my right)'; the relative frame is a ternary relationship between a located object, a reference object, and a third party who provides a viewpoint (e.g., 'the cup is to the right of the saucer', where 'right' could be either to the right of the speaker or of the addressee); the absolute frame relates object location to fixed positions in the environment (e.g., cardinal directions such as North–South or environmental features such as uphill–downhill).
Linguistic relativity: the idea that the language one speaks influences one's (non-linguistic) cognition and/or perception, also often referred to as the Sapir–Whorf hypothesis following the pioneering writings of Edward Sapir and Benjamin Lee Whorf.
Near-space visual neglect: visuospatial neglect is a disorder characterized by a deficit in the ability to orient attention toward the contralesional (usually the left) side of space. Near-space neglect is a subclass of neglect where it is restricted to near-space (usually associated with a lesion/lesions in the dorsal visual pathway, such as right posterior parietal cortex).

Box 1. The evolution of spatial communication

The origins of human language are controversial, but there is evidence that demonstratives played a central role in language evolution. Many linguists have assumed that the oldest words of human language are content words and that closed-class function words, including demonstratives, are derived from nouns and verbs by grammaticalization [121,122]. However, there are (at least) three properties that distinguish demonstratives from other closed-class items. First, although demonstratives are commonly analyzed as pronouns and determiners, they are not primarily used for grammatical purposes. In their most basic use, demonstratives occur in multimodal situations, where they are frequently accompanied by deictic pointing, a communicative device that is closely associated with the origins of language evolution [1]. Second, unlike most other closed-class items, demonstratives are universal [34,41]; many languages do not have articles, auxiliaries, and adpositions, but demonstratives seem to exist in all languages [39–41], suggesting that they are particularly important for communication and language. Third, several studies have noted that the deictic roots of demonstratives cannot (usually) be traced back to content words [39–41,123], suggesting that demonstratives emerged early in language evolution [43]. However, a recent study has questioned this view, arguing that demonstratives are commonly derived from motion verbs [122]. The main evidence for this hypothesis comes from some African languages in which a verb meaning 'to go' is phonetically similar to one of the demonstratives. There are no diachronic records of these languages, but [122] reconstruct a diachronic path from 'to go' to 'that' based on semantic and phonological considerations that is difficult to evaluate by non-experts. There is no reason to exclude the possibility that a motion verb may develop into a demonstrative, but if this has ever happened, it is a rare phenomenon. For the vast majority of languages there is no evidence that demonstratives are based on content words [39–41]. Moreover, demonstratives are not only very old but also play a crucial role in the diachronic development of grammar [34,35,40]. Across languages, we find that some of the most important grammatical markers, such as definite articles, third-person pronouns, conjunctions, copulas, and relative pronouns, are derived from demonstratives (Figure I). Because these developments are very frequent, it is now widely accepted that grammatical function words evolved from two main sources: (i) nouns and verbs and (ii) demonstratives.

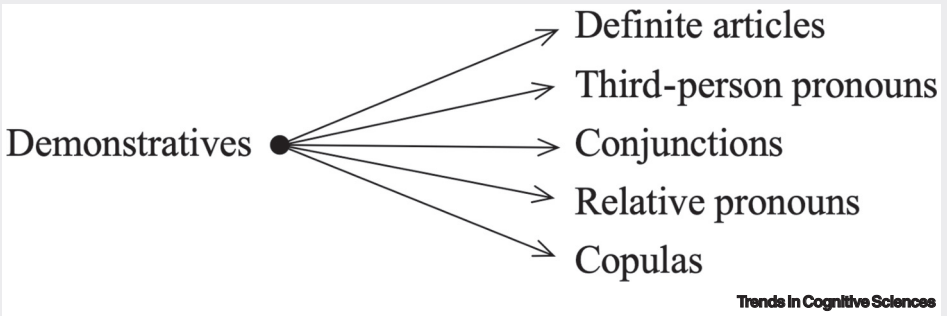

Figure I. Some frequent cross-linguistic paths of grammaticalization originating from demonstratives.

Ontogenetic ritualization: the process by which a child moves from reaching/not reaching to pointing behavior and ultimately to symbolic communication form gestures.
Person-centred: these demonstrative systems differentiate how specific demonstrative forms are applied as a function of the relative positions of speaker and hearer, and in relation to where the coordinate system is anchored. For example, various accounts of the Japanese demonstrative system distinguish between 'near a speaker', 'near an addressee', and 'far from both speaker and addressee'.
Reference frames: coordinate systems used to compute and specify the positions of an object or objects with respect to other objects.
Spatial adpositions: terms including prepositions (and postpositions in some languages) that define where an object is located with respect to another object or place. These comprise a small number of closed-class terms in language, for example 80–100 prepositions in English.

gaze [46,47]. Some languages (purportedly) require that pointing accompanies specific demonstrative forms (e.g., Goemai [48], Kilivila [49], Yucatec [50], Waroa [51], and Tiriyó [52]). Even in languages where pointing is not obligatory, gesture frequently accompanies demonstrative use (and particularly the proximal demonstrative form) [46,53–58]. Moreover, other studies have shown that being unable to see the eye gaze of a conspecific significantly reduces the production of demonstratives during a collaborative task [59]; similarly, preventing people from pointing during a similar task also reduces demonstrative production [55]. Taken together, there are compelling reasons to argue that demonstratives are fundamental terms for exploring the interplay between language and cognition across languages.

Although spatial demonstratives are universal, there is considerable cross-linguistic variation in the way demonstratives encode reference and space, as revealed by systematic 'ethnographic' studies across languages (e.g., [60]). To begin with, ~50% of the 7000 or so languages of the world (www.ethnologue.com) have binary nominal and/or adverbial systems, as in English ('this/that'; 'here/there'), whereas ~40% have three terms (as in Spanish), and the remaining languages have four or more [39,61]. Both within and between languages, a wide range of variables have been proposed to underpin demonstrative use (see [42,62] for recent reviews). Traditional

Box 2. The acquisition of spatial communication

There is little systematic research on the acquisition of demonstratives, but a recent cross-linguistic study showed that demonstratives are generally among the earliest and most frequent words children use [45]. Analyzing extensive corpus data of preschool children learning a variety of languages (including English, French, Spanish, Japanese, and Chinese), [45] found that demonstratives typically appear at around the first birthday in multimodal settings involving body-based strategies of deictic communication such as pointing, showing, and grasping. Disregarding response particles (such as English 'yes' and 'no'), demonstratives are by far the most frequent words that 1-year-old children use in all languages of the sample, accounting for ~7–11% of all child words at this young age. Demonstratives are also common in the ambient language, but children use them more frequently than their parents, suggesting that demonstratives are especially useful for young children. Other spatial expressions appear only later and are much less common. Spatial adpositions, for example, are only rarely used by 1-year-old children learning French, Spanish, Japanese, or Chinese (Figure I). English-speaking children use spatial adpositions more frequently, but compared to demonstratives they are also infrequent in English (Figure I). Interestingly, as children become older, the proportion of demonstratives decreases whereas spatial adpositions become more frequent. At the age of 5 years, spatial adpositions and demonstratives are about equally common, but, although demonstratives decrease in frequency with age, their overall proportions remain high during preschool years, compared to the proportions of demonstratives in adult language. In particular, in (adult) written language, demonstratives are rare, accounting for <1% in almost all languages of the sample. Taken together, these findings suggest a developmental shift in the conceptualization and encoding of space during the preschool years: at the beginning of first language acquisition, children use action- and body-oriented strategies of deictic communication, but when children become older they use more abstract and disembodied means to express reference and space.

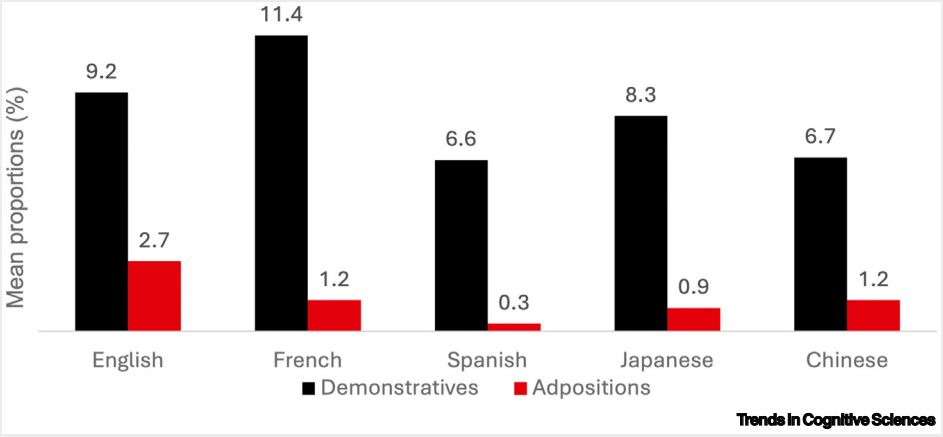

Figure I. Demonstratives and spatial adpositions in child development. The graph shows the mean proportions of demonstratives and spatial adpositions in the speech of 92 children aged 12–25 months learning English (N = 50), French (N = 18), Spanish (N = 9), Japanese (N = 9), or Chinese (N = 6) (based on corpus data analyzed in [45] including a total of 522 601 child words).

accounts of two-term demonstrative systems often assume a distance contrast in which a proximal term is used for an object (referent) near to the speaker and a distal term is used for an object far from the speaker [37,63,64].

Another important variable that has been proposed in typological studies includes the relative positions of speaker and addressee (e.g., Spanish [65–67] and Japanese [68]), and is often associated with a distinction between **person-centered** and non-person-centered demonstrative systems. About one-quarter of the demonstrative systems across the world are thought to be person-centered (M. Breunesse, PhD thesis, University of Jena, 2019), and map space onto the territories of the speaker and/or the addressee (compared to non-person-centered systems, which recognize the egocentric space of the speaker only). For instance, some accounts of the three-term Japanese demonstrative system [68] propose that one demonstrative maps onto the space near the speaker, another onto the space near the addressee, and a third relates to space far away from both speaker and addressee (Table 1).

Table 1. Demonstrative systems and linguistic diversity[a]

Two-term system			Three-term systems					
English			Spanish (non-person-centered)			Japanese (person-centered)		
	NOM	ADV		NOM	ADV		NOM	ADV
Proximal	*this*	*here*	Proximal	*este*	*aquí*	Near S	*kore*	*koko*
Distal	*that*	*there*	Medial	*ese*	*ahí*	Near H	*sore*	*soko*
			Distal	*aquel*	*allí*	Distal	*are*	*asoko*

[a]Abbreviations: ADV, adverbial demonstrative; H, hearer/addressee; NOM, nominal demonstrative; S, speaker.

Further variables thought to be central to demonstrative use in specific languages include the gaze direction of the addressee (e.g., Turkish [67,69]), the elevation of the object in the environment (e.g., Jahai [70]), object visibility (e.g., Sinhalese [71]), and object ownership (e.g., Supyire [39]). More broadly, the extent to which demonstratives should be thought of as spatial terms versus social and interactive terms has recently been a focus of attention ([42] for review), and some argue that demonstratives should not be considered as 'spatial' terms and are instead social and interactive vehicles for communication [54,67,72–74].

The diversity documented in spatial demonstrative systems across languages [60] is consistent with the extensive diversity that has been discussed for other spatial terms [25–27]. At first, one might think that this precludes the development of a unified account of spatial communication systems across languages. However, as we next review, recent empirical evidence regarding demonstrative use across languages points in the direction of such a unified account, and places action as the fundamental building block of demonstrative systems across languages.

Action as a universal constraint on spatial communication across languages

Much of the linguistic literature on demonstratives has focused on object distance as a constraint on demonstrative choice (e.g., [63]). More recently it was proposed that action rather than spatial location is a driver of the structuring of demonstrative systems across languages. Returning to the claim that space structures languages, one distinction the brain makes with respect to 'egocentric' spatial processing is between near ('peripersonal') space and far ('extrapersonal') space, served by different underlying brain systems (Figure 1) and involving action selection [75–79]. Crucially, this distinction is not only about distance but is associated with being able or unable to act on an object (implicating the presence/absence of physical control [80]). For example, it is well established that some patients with damage to their right hemisphere (usually following stroke) experience **near-space visual neglect**. When asked to point at the middle of a line (line bisection task), these patients perform poorly when the line is presented within reach, but not when it is out of reach [81,82]. Conversely, some other patients exhibit far-space neglect, with the opposite pattern of performance [83]. Moreover, there is also evidence for a shared representation of peripersonal space for oneself and another person, and a subset of neurons discovered in the left ventral premotor cortex exhibit 'mirror-like' properties with responses when an object is placed near to a person's own hand or the hand of a conspecific [84].

The idea that demonstratives may map onto space and/or action has been challenged by the observation that demonstratives can be used contrastively and flexibly. For example, one can use 'this' and 'that' to refer to objects that are both out of reach (e.g., 'this planet and that planet') or that are both within reach ('this cup and that cup') [85,86]. In addition, languages also have a 'default' or 'neutral' demonstrative that can be used regardless of the space the referent

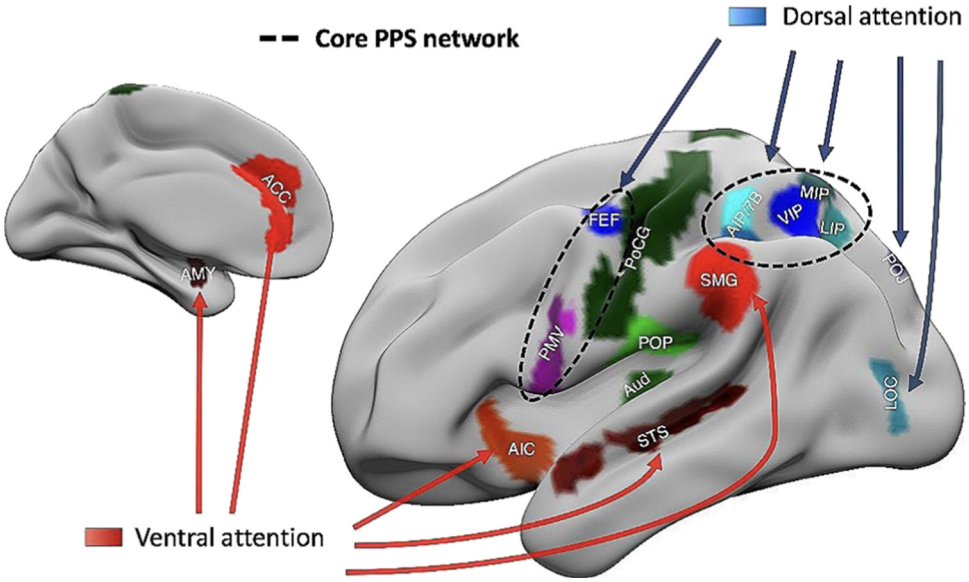

Figure 1. Graphical representation of the core and extended peripersonal space (PPS) networks. A body of neuroimaging studies in both humans and primates identified a cluster of brain regions involved in processing PPS. The figure shows a graphical representation of the core and extended PPS networks. Core areas are circles with broken lines. Regions are classified according to their relevance to large-scale brain networks and highlighted by different colors (blue, dorsal attention network; red, ventral attention network; green, sensorimotor network). Abbreviations: AAC, anterior cingulate cortex; AIC, anterior insular cortex; AIP, anterior intraparietal; AMY, amygdala; Aud, auditory belt; FEF, frontal eye field; LOC, lateral occipital cortex; PMV, ventral premotor cortex; PoCG, postcentral gyrus; POJ, parietal–occipital junction; POP, arietal operculum; SMG, supramarginal gyrus; STS, superior temporal sulcus. Figure reproduced from [79].

occupies [45,47]. Notwithstanding, several experimental studies using 'the memory game method' (Figure 2) have recently tested the mapping between reachability and demonstrative use across a wide range of languages [33,66,87–90]. For example, one study [90] manipulated both the distance an object was placed from a speaker and the position of the addressee – either spatially aligned with the speaker (sitting beside the speaker) or opposite and facing the speaker. This afforded a test of the extent to which speakers of different languages use demonstratives as a function of egocentric distance or alternatively as a function of the position of an addressee. The sample in [90] comprised 29 languages across language families, and included languages with two, three, or more demonstrative forms, those that have been argued to be person-centered (e.g., Japanese), and those that employ the absolute reference frame for other spatial terms (e.g., Tzeltal). Of note was that this study included sufficient numbers of speakers of each language to be able to separate out variation within a language and variation between languages (compared to earlier field studies with very small numbers of participants; e.g., [60]). The results showed that all languages tested have a proximal demonstrative that is used most when the object being referred to is reachable by the speaker, and a second demonstrative form that maps onto space out of reach of the speaker from an egocentric perspective (Figure 3, Key figure). Moreover, a range of languages with three-term demonstrative systems (e.g., Japanese, Korean, Finnish, Lithuanian, and Georgian) have a third demonstrative form which indicates that the referent is reachable by the addressee, thus extending action to the actions that can be performed by a conspecific when the speaker is not in a position to act on the object [80,91].

Further evidence for this basic mapping between peripersonal/extrapersonal space and demonstrative forms comes from studies that manipulated reachability while keeping object distance

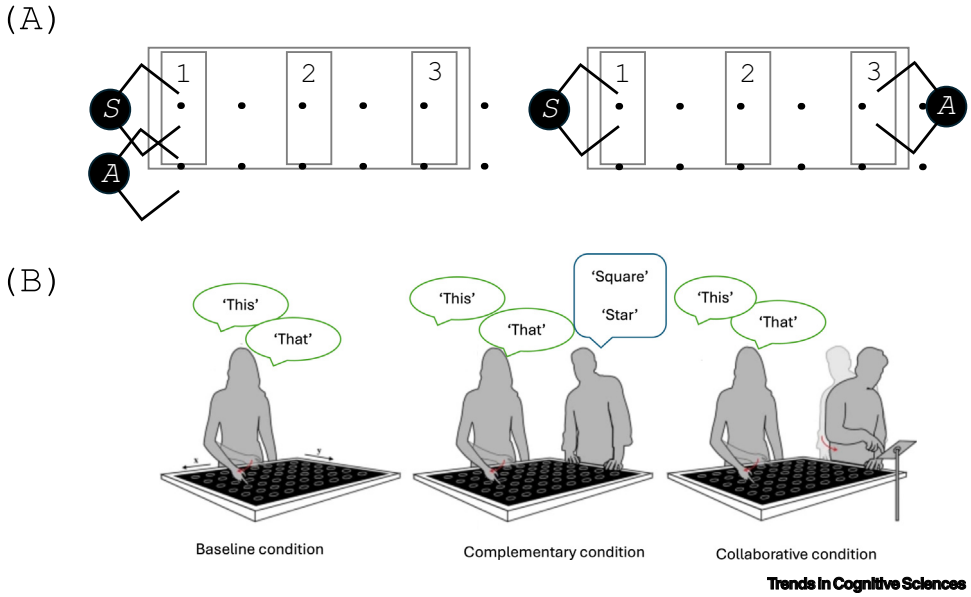

Figure 2. Testing spatial communication using the 'memory game' paradigm. In the 'memory game' paradigm participants think they are taking part in a memory study and that they are in the 'language condition'. Once an object has been placed, the participant points at the object and describes it using a combination of a demonstrative, a color, and a shape (e.g., 'this/that red triangle'), with memory probes used throughout the experiment to maintain the guise of the memory experiment. Studies using this method have been conducted across many languages and have tested a range of variables, with all studies revealing that reachability influences demonstrative choice. (A) The spatial arrangements in [90] where the addressee (*A*) is positioned either beside the speaker/participant (*S*) (thus sharing the same spatial perspective, left picture) or opposite the speaker/participant (right picture). Adapted from [90]. (B) The conditions used in a variant of the method in [99]. The participant described the target shapes in the absence of a conspecific (baseline condition, left picture), in the presence of a conspecific who independently named the target shapes after the participant had described them (complementary condition, middle picture), or in the presence of a conspecific whose naming was dependent on the descriptions produced by the participant (collaborative condition, right picture). The collaborative condition with codependency between the behaviors of the participants led to an extension of the proximal term to the reachable space of the addressee in Danish.

constant. In one study reachability was manipulated by asking participants to point at an object (placed at different distances) with either their arm/hand or with a stick [66]. Mirroring findings showing an extension of **near-space neglect** to far-space when patients point with a stick (e.g., [81–83]), the results showed an extension of the use of the proximal terms in English ('this') and Spanish (*este*) to object locations outside of arm's reach but reachable by the stick when pointing with the stick. In another study [87], objects were placed in the sagittal and lateral planes in front of participants, and they pointed at the objects with either their left or right hand. In doing so, there were equidistant locations on the left and right of the participants that were reachable when pointing with one hand but not the other. 'This' was used more frequently to refer to an object when pointing with a hand that could reach it compared to pointing with the other hand (irrespective of the handedness of the participant).

From a neural perspective, there is evidence that spatial language processing goes beyond the so-called 'default' language network that has been implicated in language processing across languages (e.g., [92]). Imaging studies during the comprehension of spatial demonstratives [93] and spatial adpositions [94–96], as well as neuropsychological studies on adpositions (e.g., [97]), implicate an extended brain network involving supramarginal gyrus, angular gyrus, and precuneus that is associated with visuospatial processing (consistent with the activation of a **dorsal 'where' stream**) (Figure 1; see [98] for a review).

Key figure

Testing demonstrative choice across languages

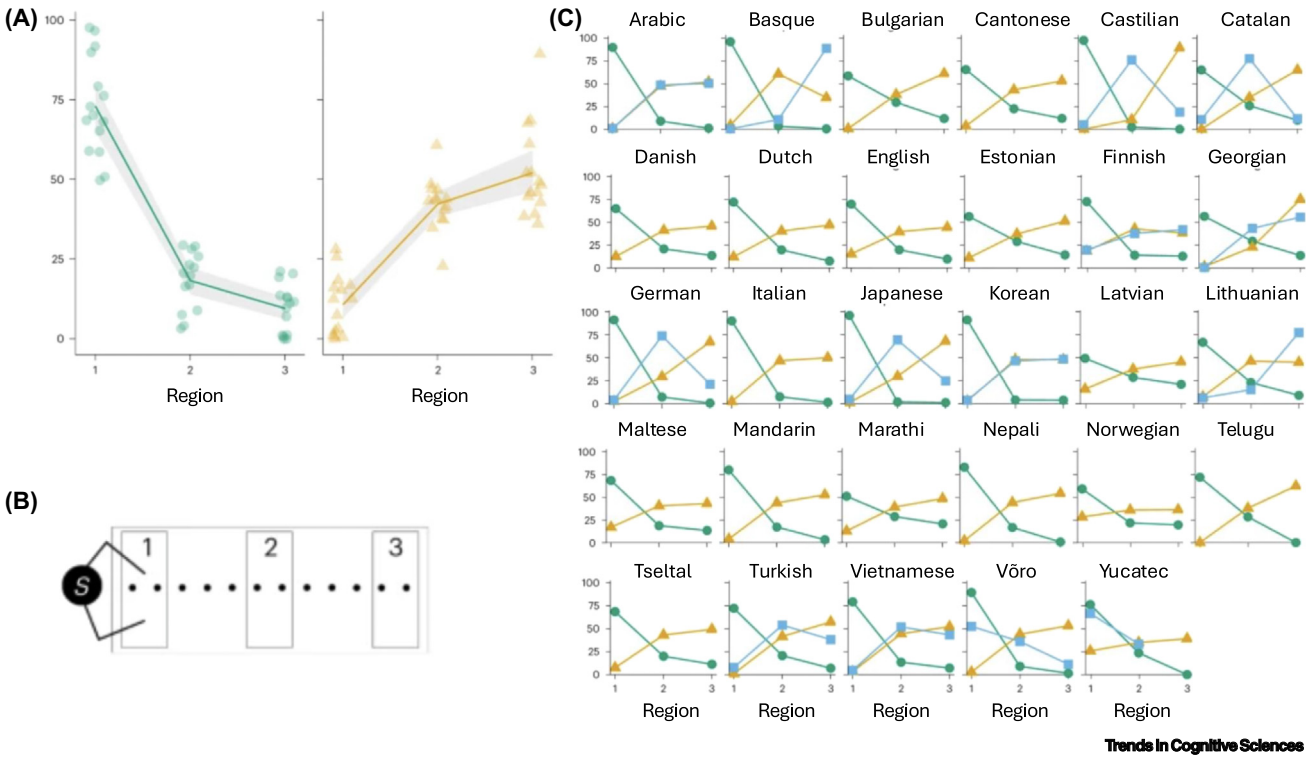

Figure 3. Effects of reachability across 29 languages where all languages possess a 'proximal' demonstrative (green) that maps onto reachable space and a 'distal' term (orange) that maps onto non-reachable space. (A) The normalized frequency distributions across languages of proximal (left) and distal (right) terms. (B) The key to regions of space. (C) Reachability effects for each of the 29 languages tested in [90]. Adapted from [90].

The extension of reachability effects to an addressee, although 'explicit' in some languages with three-term demonstrative systems, is also important for speakers of languages with only binary systems. One study [99] manipulated the extent to which face-to-face Danish-speaking participants interacted during a task that involved speakers pointing and using demonstratives to refer to objects at various locations on a table (in a variant of the memory game; Figure 2). In one (baseline) condition the participants completed the task alone, but in two other conditions a confederate was standing at 90° to the participant at another side of the table on which the object location grid was placed. In one condition the confederate named the target shapes independently of the task performed by the (speaker) participant, and in another (collaborative) condition the tasks of the confederate and speaker were codependent (and hence collaborative). It was found that participants in the context of collaborative interaction remapped their action space and used the proximal demonstrative when the object being referred to was reachable by the collaborative partner.

Evidence of taking someone else's perspective in spatial communication as a function of action has also been found with the so-called 'projective' spatial adpositions such as 'to the left of' and 'to the right of' [16,100]. For example, in one study [16] participants described the positions

of two objects placed side-by-side on a table in a photograph in which a person was shown (behind the table, facing the participant) either reaching or not reaching toward one of the objects. When the person in the picture was reaching, participants were more likely to describe the objects using left/right from the perspective of the person in the picture rather than from their own (egocentric) perspective.

These studies show that action plays a key role in determining whether one chooses to describe space from one's own perspective or from the perspective of another. Moreover, for all 29 languages in [90], significant within-language variation was found, indicating that speakers of all languages do not use the demonstratives available to them in their language uniformly. This suggests that, even for languages such as Japanese that have a specific demonstrative that can be used to refer to the reachable space of the addressee, some Japanese speakers still choose to describe the object as far from themselves as opposed to reachable by the addressee. Following the study of Rocca *et al.* [99], one might expect that increasing the interaction between Japanese participants during a task might also increase the likelihood of speakers choosing the perspective of the addressee.

The extension of reachability to an addressee places perspective-taking at the heart of spatial communication. Recent work suggests that adults automatically and effortlessly monitor someone else's perspective during perceptual decision-making [101–103]. Moreover, the processing of space itself is affected by the presence and relative position of a conspecific during a task [104–106], suggesting that the processing of space is partly determined by social factors. Consistent with this, the only neuroimaging study to date that examined the comprehension of spatial demonstratives [93] showed that processing of demonstratives also includes activation of frontal regions, including frontal eye fields, that have been implicated in reference frame shifting and attention reorienting [107,108].

Spatial demonstratives: uniting action, attention, and theory of mind

Spatial communication systems may be regarded as fundamental to language and language development because they serve several interlinked functions that are crucial for communication. First, spatial communication in the form of demonstratives is directly linked to action, and specifically to the distinction between reachable and non-reachable space. Furthermore, this possible universal feature of the demonstrative systems across the world is also explicitly or implicitly extended in languages to refer to reachability from the perspective of someone else who can reach an object that the speaker is unable to reach. Thus, the second key feature of spatial communication systems that is common across demonstratives (and adpositions) is the ability to take a perspective different from one's own. In turn, this requires the attention of an addressee to be monitored so as to anticipate how they might act and interact. Spatial communication therefore forms a natural bridge between perception, action, and theory of mind.

It has long been recognized that demonstratives serve to create and manipulate joint attention in face-to-face communication [34,109]. Recent analyses of attention immediately preceding the use of demonstratives by infants and caregivers in naturalistic interactions have shown that such communication is usually preceded by coordinated attention to the space in which the referent is situated, and attention is subsequently directed to the intended reference within the attended space via a combination of demonstratives, eye gaze, and gesture [58].

Several recent treatments of demonstratives have also examined how demonstratives are used as a function of disjoint attention between the speaker and addressee [67,69,110], where demonstratives are used to draw the attention of an addressee back to the interaction at hand. In

a series of online studies, one recent report [67] manipulated the relative positions of the speaker and addressee and the gaze direction of the addressee, such that the addressee was looking at the same or a different object in an array from the speaker. When speaker and addressee gaze was misaligned, Turkish speakers thought the speaker would use the (middle term) *şu* more frequently compared to the scenes where their gaze was aligned, supporting the view that Turkish has a specific term used to align speaker and addressee gaze onto the intended referent object. Whether demonstrative systems in languages that lack such an overt term also serve this function remains to be established.

The early appearance of demonstratives in language acquisition may presage the importance of these terms in the development of perspective taking, social attention, and theory of mind [34,43,111], paving the way for the development of other forms of spatial communication and of communication more broadly. The emergence of demonstratives building on deictic pointing, the universality of demonstrative forms, and the (likely) universal action distinctions that demonstrative systems make may also give clues to how language itself developed (*cf* [112]). Accounts of language evolution include the idea that language evolved from gesture [1,113–115]. We can speculate that the universal reachable/non-reachable distinction may be a gateway to the development of perspective taking in language (consistent with **ontogenetic ritualization** [1,115]). Moreover, demonstratives may offer a possible clue to unifying the gestural theory of the origins of communication with alternative accounts maintaining that language evolved to serve social functions [116,117] as well as spatial navigational needs [10,114,118].

Concluding remarks

In this review we have shown that spatial demonstratives offer a crucial window into the relationship between language and cognition. In contrast to the view that spatial communications systems are an exemplar of linguistic diversity [22,25–31], recent evidence suggests that inherently multimodal spatial demonstrative systems are built from a universal distinction that places action at the center of spatial communication. This reinvigorates the debate regarding the centrality of egocentric space, while also showing that these fundamental terms embody a confluence of action, interaction, attention, and theory of mind. Although developmental studies suggest that the acquisition of spatial demonstratives is protracted (e.g., [69,119]), future studies will be necessary to understand exactly how the use of demonstratives unfolds in development, whether (in addition to action) there are other common constraints on demonstrative systems across languages (e.g., [32,33,120] for discussion), and how brain mechanisms support demonstrative use across languages (see Outstanding questions).

Acknowledgments

This research was supported by EU Horizon 2020 Innovative Training Network (ITN) Marie Skłodowska-Curie Action grant 676063 (DCOMM) to K.R.C. and colleagues.

Declaration of interests

No interests are declared.

References

1. Tomasello, M. (2008) *The Origins of Human Communication*, The MIT Press
2. Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*, University of Chicago Press
3. Casasanto, D. and Boroditsky, L. (2008) Time in mind: using space to think about time. *Cognition* 106, 379–593
4. Talmy, L. (2018) *The Targeting System of Language*, MIT Press
5. Langacker, R. (1987) *Foundations of Cognitive Grammar. Vol I: Theoretical Prerequisites*, Stanford University Press
6. Lakoff, G. (1987) *Women, Fire and Dangerous Things*, University of Chicago Press
7. Talmy, L. (2000) *Toward a Cognitive Semantics (Vols 1 and 2)*, MIT Press
8. Goldberg, A. (2006) *Constructions at Work: The Nature of Generalization in Language*, Oxford University Press
9. Kövecses, S. (2020) *Extended Conceptual Metaphor Theory*, Cambridge University Press
10. O'Keefe, J. and Nadel, L. (1978) *The Hippocampus as a Cognitive Map*, Oxford University Press
11. Tversky, B. (2023) *Mind in Motion*, Bradford Books
12. Fodor, J. (2009) *LOT2. The Language of Thought Revisited*, Oxford University Press

Outstanding questions

In addition to the reachable/non-reachable distinction, are there any other universal constraints on demonstrative systems? There is evidence that some other constraints, such as visibility, elevation, and attention, affect demonstrative use in some languages, including languages that lack specific demonstratives forms that mark the contrast. Systematic studies across languages will be necessary to assess whether these variables universally affect the use of demonstratives.

What is the relationship between the reachable/non-reachable distinction and other (non- reachability-based) demonstratives within and between languages? Demonstratives can be used flexibly, but whether and how such flexibility relates to a basic action distinction has not yet been established.

What factors determine how children use demonstratives at different points in development? There is some evidence in a limited range of languages that the use of demonstratives, like the use of other spatial terms, continues to develop over a protracted period. However, systematic studies to date have yet to be conducted regarding how demonstratives are acquired and how their use changes over time.

What brain systems underpin the use of demonstratives across languages? Only one study to date has examined the brain systems that underpin the comprehension of demonstratives in heard narratives. Evidence for a mapping between peripersonal and extrapersonal space processing and specific demonstrative forms has yet to be forthcoming.

What is the relationship between spatial demonstratives and non-spatial uses of demonstratives across languages? Demonstratives, like other spatial terms, can be also used non-spatially, such as in temporal expressions ('that party', 'this day') and can also be used as grammatical markers in text. Although there are some theoretical treatments of the relationship between spatial and non-spatial uses of demonstratives, future empirical studies are much needed in this area.

13. Pinker, S. (1994) *The Language Instinct*, HarperCollins
14. Li, P. and Gleitman, L. (2002) Turning the tables: language and spatial reasoning. *Cognition* 83, 265–294
15. Kant, I. (1768/1991) On dem ersten Grunde des Unterschieds der Gegenden im Raume [On the first ground of the distinction of regions in space]. In *The Philosophy of Right and Left: Incongruent Counterparts and the Nature of Space* (van Cleve, J. and Frederick, R.E., eds), pp. 27–33, Kluwer
16. Tversky, B. and Hard, B.M. (2009) Embodied and disembodied cognition: spatial perspective-taking. *Cognition* 110, 124–129
17. Piaget, J. and Inhelder, B. (1956) *The Child's Conception of Space*, Routledge & Kegan Paul
18. Gollege, R. (1992) Place recognition and wayfinding: making sense of space. *Geoforum* 23, 199–214
19. Shelton, A.L. and McNamara, T.P. (1997) Multiple views of spatial memory. *Psychon. Bull. Rev.* 4, 102–106
20. Whorf, B.L. (1956) *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*, MIT Press
21. Wolff, P.M. and Holmes, K.J. (2011) Linguistic relativity. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 253–265
22. Evans, N. and Levinson, S.C. (2009) The myth of language universals: language diversity and its importance for cognitive science. *Behav. Brain Sci.* 32, 429–448
23. Landau, B. and Jackendoff, R. (1993) 'What' and 'where' in spatial language and spatial cognition. *Behav. Brain Sci.* 16, 217–265
24. Coventry, K.R. and Garrod, S.C. (2004) *Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions*, Taylor Francis
25. Pederson, E. *et al.* (1998) Semantic typology and spatial conceptualization. *Language* 74, 557–589
26. Levinson, S.C. (2003) *Space in Language and Cognition: Explorations in Cognitive Diversity*, Cambridge University Press
27. Majid, A. *et al.* (2004) Can language restructure cognition? The case for space. *Trends Cogn. Sci.* 8, 108–114
28. Levinson, S.C. (1996) Frames of reference and Molyneux's question: cross-linguistic evidence. In *Language and Space* (Bloom, P. *et al.*, eds), pp. 109–169, MIT Press
29. Haun, D.B.M. and Rapold, C.J. (2009) Variation in memory for body movements across cultures. *Curr. Biol.* 19, R1068-R109
30. Haun, D.B.M. *et al.* (2009) Cognitive cladistics and cultural override in Hominid spatial cognition. *Proc. Natl. Acad. Sci. U. S. A.* 103, 17568–17573
31. Haun, D.B.M. *et al.* (2011) Plasticity of human spatial memory: spatial language and cognition covary across cultures. *Cognition* 119, 70–80
32. Li, P. *et al.* (2011) Spatial reasoning in Tseltal Mayan. *Cognition* 120, 33–53
33. Coventry, K.R. *et al.* (2014) Spatial demonstratives and perceptual space: describing and remembering object location. *Cogn. Psychol.* 69, 46–70
34. Diessel, H. (2006) Demonstratives, joint attention, and the emergence of grammar. *Cogn. Linguist.* 17, 463–489
35. Liszkowski, U. *et al.* (2012) A prelinguistic gestural universal of human communication. *Cogn. Sci.* 36, 698–713
36. Lüke, C. *et al.* (2020) Integrated communication system: gesture and language acquisition in typically developing children and children with LD and DLD. *Front. Psychol.* 11, 118
37. Bühler, K. (1934) *Sprachtheorie*, Fischer
38. Levinson, S.C. (2004) Deixis and pragmatics. In *The Handbook of Pragmatics* (Horn, L. and Ward, G., eds), pp. 97–121, Blackwell
39. Diessel, H. (1999) *Demonstratives. Form, Function, and Grammaticalization*, John Benjamins
40. Himmelman, N. (1996) *Deiktikon, Artikel, Nominalphrase. Zur Emergenz syntaktischer Struktur*, Narr
41. Dixon, R.M.W. (2003) Demonstratives: a cross-linguistic typology. *Stud. Lang.* 27, 61–112
42. Diessel, H. and Coventry, K.R. (2020) Demonstratives in spatial language and social interaction: an interdisciplinary review. *Front. Psychol.* 11, 555265
43. Diessel, H. (2013) Where does language come from? Some reflections on the role of deictic gesture and demonstratives in the evolution of language. *Lang. Cogn.* 5, 239–249
44. Clark, E.V. (1978) From gesture to word: on the natural history of deixis in language acquisition. In *Human Growth and Development* (Bruner, J.S. and Garton, A., eds), pp. 85–120, Oxford University Press
45. Diessel, H. and Monakhov, S. (2023) Acquisition of demonstratives in cross-linguistic perspective. *J. Child Lang.* 50, 922–953
46. Enfield, N.J. (2003) Demonstratives in space and interaction: data from Lao speakers and implications for semantic analysis. *Language* 79, 82–117
47. Levinson, S.C. (2018) Demonstratives – patterns of diversity. In *Demonstratives in Cross-Linguistic Perspective* (Levinson, S.C. *et al.*, eds), pp. 1–42, Cambridge University Press
48. Hellwig, B. (2003) *The Grammatical Coding of Postural Semantics in Goemai (A West Chadic Language of Nigeria)*, Ponsen & Looijen
49. Senft, G. (2004) Aspects of spatial deixis in Kilivila. In *Deixis and Demonstratives in Oceanic Languages* (Senft, G., ed.), pp. 59–80, Australian National University
50. Bohnemeyer, J. (2018) Yucatec demonstratives in interaction: spontaneous versus elicited data. In *Demonstratives in Cross-linguistic Perspective* (Levinson, S.C. *et al.*, eds), pp. 176–205, Cambridge University Press
51. Herrmann, S. (2018) Waroa demonstratives. In *Demonstratives in Cross-linguistic Perspective* (Levinson, S.C. *et al.*, eds), pp. 282–302, Cambridge University Press
52. Meira, S. (2018) Tiriyó – non-contrastive exophoric uses of demonstratives. In *Demonstratives in Cross-Linguistic Perspective* (Levinson, S.C. *et al.*, eds), pp. 222–241, Cambridge University Press
53. Stukenbrock, A. (2015) *Deixis in der face-to-face-Interaktion*, Mouton de Gruyter
54. Gipper, S. (2017) Pre-semantic pragmatics encoded: a non-spatial account of Yurakaré demonstratives. *J. Pragmat.* 120, 122–143
55. Bangerter, A. (2004) Using pointing and describing to achieve joint focus on attention in dialogue. *Psychol. Sci.* 15, 415–419
56. Piwek, P. *et al.* (2008) Proximal and distal language and cognition: evidence from deictic demonstratives in Dutch. *J. Pragmat.* 40, 694–718
57. Cooperrider, K. (2016) The co-organization of demonstratives and pointing gestures. *Discourse Process.* 53, 632–656
58. Todisco, E. *et al.* (2021) The temporal dynamics of deictic communication. *First Lang.* 41, 154–178
59. Garciá, J.L.P. *et al.* (2017) Bodily constraints contributing to multimodal referentiality in humans: the contribution of a depigmented sclera to proto-declaratives. *Lang. Commun.* 54, 73–81
60. Levinson, S.C. *et al.*, eds (2018) *Demonstratives in Cross-Linguistic Perspective*, Cambridge University Press
61. Diessel, H. (2013) Distance contrasts in demonstratives. In *World Atlas of Language Structures Online* (Dryer, M.S. and Haspelmath, M., eds), Zenodo
62. Peeters, D. *et al.* (2021) A conceptual framework for the study of demonstrative reference. *Psychon. Bull. Rev.* 28, 409–433
63. Anderson, S.R. and Keenan, E.L. (1985) Deixis. In *Language Typology and Syntactic Description III* (Shopen, T., ed.), pp. 259–308, Cambridge University Press
64. Diessel, H. (2014) Demonstratives, frames of reference, and semantic universals of space. *Lang. Ling. Compass* 8, 116–132
65. Jungbluth, K. (2003) Deictics in the conversational dyad. In *Deictic Conceptualizations of Space, Time and Person* (Lenz, F., ed.), pp. 13–40, John Benjamins
66. Coventry, K.R. *et al.* (2008) Language within your reach: near–far perceptual space and spatial demonstratives. *Cognition* 108, 889–895
67. Rubio-Fernandez, P. (2022) Demonstrative systems: from linguistic typology to social cognition. *Cogn. Psychol.* 139, 101519
68. Hasegawa, Y. (2012) Deictic and anaphoric uses of the Japanese demonstratives ko-so-a. *J. Japan Linguist.* 28, 44–61
69. Küntay, A. and Özyürek, A. (2006) Learning to use demonstratives in conversation: what do language specific strategies in Turkish reveal? *J. Child Lang.* 33, 303–320
70. Forker, D. (2020) Elevation as a grammatical and semantic category of demonstratives. *Front. Psychol.* 11, 1712
71. Chandralal, D. (2010) *Sinhala*, John Benjamins
72. Hanks, W.F. (1990) *Referential Practice*, Chicago University Press

73. Peeters, D. and Özyürek, A. (2016) This and that revisited: a social and multimodal approach to spatial demonstratives. *Front. Psychol.* 7, 222
74. Peeters, D. *et al.* (2015) Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition* 136, 64–84
75. Legrand, D. *et al.* (2007) Close to me: multisensory space representations for action and pre-reflexive consciousness of oneself-in-the-world. *Conscious. Cogn.* 16, 687–699
76. Iriki, A. *et al.* (1996) Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 7, 2325–2330
77. Bufacchi, R.J. and Iannetti, G.D. (2018) An action field theory of peripersonal space. *Trends Cogn. Sci.* 22, 1076–1090
78. Serino, A. (2019) Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neurosci. Biobehav. Rev.* 99, 138–159
79. Basile, G.A. *et al.* (2024) Neuroanatomical correlates of peripersonal space: bridging the gap between perception, action, emotion and social cognition. *Brain Struct. Funct.* 229, 1047–1072
80. Brovold, A. and Grush, R. (2012) Towards an (improved) interdisciplinary investigation of demonstrative reference. In *Perception, Realism and the Problem of Reference* (Athanassios, A. and Machamer, P., eds), pp. 11–42, Cambridge University Press
81. Berti, A. and Frassinetti, F. (2000) When far becomes near: remapping of space by tool use. *J. Cogn. Neurosci.* 12, 415–420
82. Longo, M.R. and Lourenco, S.F. (2006) On the nature of near space: effects of tool use and the transition to far space. *Neuropsychologia* 44, 977–981
83. Neppi-Mòdona, M. *et al.* (2007) Bisecting lines with different tools in right brain damaged patients: the role of action programming and sensory feedback in modulating spatial remapping. *Cortex* 43, 397–410
84. Brozzoli, C. *et al.* (2013) A shared representation of the space near oneself and others in the human premotor cortex. *Curr. Biol.* 23, 1764–1768
85. Kemmerer, D. (1999) Near and far in language and perception. *Cognition* 73, 35–63
86. Bonfiglioli, C. *et al.* (2009) A kinematic approach to the conceptual representations of this and that. *Cognition* 111, 270–274
87. Caldano, M. and Coventry, K.R. (2019) Spatial demonstratives and perceptual space: to reach or not to reach? *Cognition* 191, 103989
88. Reile, M. *et al.* (2020) Demonstratives as spatial deictics or something more? Evidence from Common Estonian and Võro. *Folia Linguistica* 54, 167–195
89. Skilton, A. and Peeters, D. (2021) Cross-linguistic differences in demonstrative systems: comparing spatial and non-spatial influences on demonstrative use in Ticuna and Dutch. *J. Pragmat.* 180, 48–265
90. Coventry, K.R. *et al.* (2023) Spatial communication systems across language reflect universal action constraints. *Nat. Hum. Behav.* 7, 2099–2110
91. Imai, S. (2009) *Spatial Deixis: How Finely Do Languages Divide Space?*, VDM
92. Malik-Moraleda, S. *et al.* (2022) An investigation across 45 languages and language families reveals a universal language network. *Nat. Neurosci.* 25, 1014–1019
93. Rocca, R. *et al.* (2020) Language beyond the language system: dorsal visuospatial pathways support processing of demonstratives and spatial language during naturalistic fast fMRI. *NeuroImage* 216, 116128
94. Noordzij, M.L. *et al.* (2008) Neural correlates of locative prepositions. *Neuropsychologia* 46, 1576–1580
95. Damasio, H. *et al.* (2001) Neural correlates of naming actions and of naming spatial relations. *NeuroImage* 13, 1053–1064
96. Coventry, K.R. *et al.* (2013) Multiple routes to mental animation: language and functional relations drive motion processing for static images. *Psychol. Sci.* 24, 1379–1388
97. Tranel, D. and Kemmerer, D. (2004) Neural correlates of locative prepositions. *Cogn. Neuropsychol.* 21, 719–749
98. Kemmerer, D. (2019) *Concepts in the Brain. The View from Cross-Linguistic Diversity*, Oxford University Press
99. Rocca, R. *et al.* (2019) This is for you: social modulations of proximal vs. distal space in collaborative interaction. *Sci. Rep.* 9, 14967
100. Tosi, A. *et al.* (2019) Speakers' use of agency and visual context spatial descriptions. *Cognition* 194, 104070
101. Samson, D. *et al.* (2010) Seeing it their way: evidence for rapid and involuntary computation of what other people see. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1255–1266
102. Ward, E. *et al.* (2019) Spontaneous vicarious perception of the content of another's visual perspective. *Curr. Biol.* 29, 874–880
103. Lukošiūnaitė, I. *et al.* (2024) The influence of another's actions and presence on perspective taking. *Sci. Rep.* 14, 4971
104. Teneggi, C. *et al.* (2013) Social modulation of peripersonal space boundaries. *Curr. Biol.* 23, 406–411
105. Pellencin, E. *et al.* (2018) Social perception of others shapes one's own multisensory peripersonal space. *Cortex* 104, 163–179
106. Galati, A. *et al.* (2019) Social and configural effects on the cognitive dynamics of perspective-taking. *J. Mem. Lang.* 104, 1–24
107. Wallentin, M. (2012) The role of the brain's frontal eye fields in constructing framer of reference. *Cogn. Process.* 13, 359–363
108. Wallentin, M. *et al.* (2006) Frontal eye fields involved in shifting frame of reference within working memory for scenes. *Neuropsychologia* 46, 399–408
109. Clark, H.H. (1996) *Using Language*, Cambridge University Press
110. Evans, N. *et al.* (2018) The grammar of engagement I: framework and initial exemplification. *Lang. Cogn.* 10, 110–140
111. Rubio-Fernandez, P. (2023) Cultural evolutionary pragmatics: investigating the codevelopment and coevolution of language and social cognition. *Psychol. Rev.* 131, 18–35
112. Slobin, D. (2005) From ontogenesis to phylogenesis: what can child language tell us about language evolution? In *Biology and Knowledge Revisited: From Neurogenesis to Psychogenesis* (Langer, S.T. and Parker, C., eds), pp. 255–285, Lawrence Erlbaum Associates
113. Arbib, M. (2012) *How the Brain Got Language: The Mirror System Hypothesis*, Oxford Univ. Press
114. Corbalis, M. (2017) Language evolution: a changing perspective. *Trends Cogn. Sci.* 21, 229–236
115. Spranger, M. and Steels, L. (2014) Discovering communication through ontogenetic ritualisation. In *4th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics*, pp. 14–19, IEEE
116. Seyfarth, R.M. and Cheney, D.L. (2014) The evolution of language from social cognition. *Curr. Opin. Neurobiol.* 28, 5–9
117. Harari, Y.N. (2014) *A Brief History of Humankind*, Harvill Secker
118. Berwick, R.C. and Chomsky, N. (2016) *Why Only Us? Language and Evolution*, MIT Press
119. Guijarro-Fuentes, P. *et al.* (2022) Children's use of demonstrative words: spatial deictics beyond infancy. *J. Child Lang.* 51, 952–964
120. Chen, S. *et al.* (2023) An information-theoretic approach to the typology of spatial demonstratives. *Cognition* 240, 105505
121. Hopper, P.J. and Traugott, E.C. (1993) *Grammaticalization*, Cambridge University Press
122. Heine, B. *et al.* (2020) Where do demonstratives come from? STUF – Language Typology and Universals. 73, 403–434
123. Traugott, E.C. (1982) From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In *Perspectives on Historical Linguistics* (Lehmann, W.P. and Malkiel, Y., eds), pp. 245–271, John Benjamins