

Review



Cite this article: Kuhlen AK, Abdel Rahman R. 2023 Beyond speaking: neurocognitive perspectives on language production in social interaction. *Phil. Trans. R. Soc. B* **378**: 20210483.
<https://doi.org/10.1098/rstb.2021.0483>

Received: 8 July 2022

Accepted: 16 December 2022

One contribution of 15 to a discussion meeting issue 'Face2face: advancing the science of social interaction'.

Subject Areas:

neuroscience, cognition, behaviour

Keywords:

language production, social interaction, dialogue, psycholinguistics, theory of mind, joint action

Author for correspondence:

Anna K. Kuhlen

e-mail: anna.kuhlen@psych.rwth-aachen.de

Beyond speaking: neurocognitive perspectives on language production in social interaction

Anna K. Kuhlen and Rasha Abdel Rahman

Department of Psychology, Humboldt-Universität zu Berlin, 12489 Berlin, Germany

AKK, 0000-0001-8834-0592; RA, 0000-0002-8438-1570

The human faculty to speak has evolved, so has been argued, for communicating with others and for engaging in social interactions. Hence the human cognitive system should be equipped to address the demands that social interaction places on the language production system. These demands include the need to coordinate speaking with listening, the need to integrate own (verbal) actions with the interlocutor's actions, and the need to adapt language flexibly to the interlocutor and the social context. In order to meet these demands, core processes of language production are supported by cognitive processes that enable interpersonal coordination and social cognition. To fully understand the cognitive architecture and its neural implementation enabling humans to speak in social interaction, our understanding of how humans produce language needs to be connected to our understanding of how humans gain insights into other people's mental states and coordinate in social interaction. This article reviews theories and neurocognitive experiments that make this connection and can contribute to advancing our understanding of speaking in social interaction.

This article is part of a discussion meeting issue 'Face2face: advancing the science of social interaction'.

1. Introduction

As humans we have the basic need to connect with each other, to share our ideas, thoughts, and feelings. Language is the central instrument to meet this need. When we speak to communicate, we typically engage in social interaction with others. Yet mainstream psycholinguistic research mainly investigates language production in settings devoid of any form of social interaction. This approach corresponds to an understanding of language as an instrument of the human mind for processing information and structuring thinking (versus as an instrument enabling social interaction, for discussion see [1,2]). Research in this tradition has provided the basis for mechanistic theories of language production, which dissect into subsystems the processing steps required for an individual speaker to get from forming a communicative intention to generating overt speech. While these theories have been instrumental for understanding the cognitive processes underlying isolated language production, it is vital to scale up to settings in which language is typically produced, namely, in social interaction (see [1–3]).

The core processes of language production are shaped by the unique demands social interaction places on our cognitive system. Moreover, the demands of social interaction are met by engaging additional cognitive processes and neural structures; those that support social cognition and interpersonal coordination. Theories of language production need to be embedded in a larger understanding of the human cognitive system and the mechanisms that support successful social (inter-)actions. Recent advancements in this direction cross disciplinary boundaries and adapt methods from experimental psychology, social psychology and cognitive neuroscience, and connect our understanding of how humans produce language to our understanding of how humans gain

insight into other people's minds and coordinate their actions in social encounters.

The goal of this article is to pull together insights and perspectives on how neurocognitive processes underlying speaking coordinate with, and are shaped by, processes supporting social cognition and social interaction. We will focus on three demands that social interaction places on the language production system: (i) the need to coordinate speaking with listening, (ii) the need to coordinate own behaviour with the conversational partner's behaviour as part of a joint action, and (iii) the need to flexibly adapt speaking to a particular conversational partner.

2. Language production in social interaction: more than speaking

Speaking in a social interaction is something we do every day and without thinking much about it. Yet, as simple as it may appear, it is a highly complex task. Interlocutors need to accomplish the sheer act of producing speech, starting with forming the intention to communicate, generating the message, selecting the right words to express this message, encoding it as grammatical and phonological information, and finally, translating this into motor commands that execute articulation ([4]; for recent review see [5]). Speaking is also commonly accompanied by hand gestures, facial displays, or other bodily cues that need to be integrated, both in time and content, with the spoken message (e.g. [6]). What is more, in a social setting, additional demands are placed upon a speaker's cognitive system that go beyond the core processes of language production. These demands are met by cognitive mechanisms and neural structures supporting social cognition and joint action coordination.

3. Integrating speaking and listening

One source of insights into the nature of the demands spoken social interactions place on speakers have been ethnomethodological studies that describe and analyse patterns of behaviour through observations of natural occurring spoken interactions (e.g. [7–9]). A central topic in this research tradition has been turn-taking, which describes the alternation of speaking turns between different speakers (e.g. [9]). This characteristic of conversational speech comes with consequences for the cognitive system and has informed and inspired experimental psychologists to investigate the cognitive mechanisms that could support such behaviour (for discussion see [10]). As we will see, the human cognitive system seems well equipped to meet this challenge.

One basic consequence that follows from the turn-taking structure of conversation is the tight integration of speaking and listening: processes of language production become interwoven with processes of language comprehension—and may happen in parallel. Crucial insights how this may be achieved have come from connecting language production to general principles of human cognition: cognitive theories of action planning and action execution have proposed that the perception of action and the execution of action share common representational structures [11]. Mirror neurons in the brain that fire both in response to viewing an action and in response

to executing an action, provide a neuroscientific basis for this mechanism (e.g. [12–14]).

In language, as well, the argument has been made that the representations underlying language comprehension and those underlying language production are of the same format (e.g. [15–18]). In fact, this principle has been proposed to be a fundamental building block of language use in social interaction [3]. According to this proposal, speaking is facilitated by processing the partner's utterance: via a simple priming mechanism, it is efficient for speakers to re-use linguistic structures (e.g. the same lexical expression or syntactic construction) previously used by their conversational partner. This principle has been offered as a mechanistic explanation for the frequently observed behaviour of interlocutors to entrain on their linguistic behaviour and become more similar to each other over time. Indeed, a large body of literature has demonstrated that conversational partners converge over the course of an interaction on different levels of representation, ranging from the phonetic level (e.g. [19]) to the use of similar speech-accompanying gestures (e.g. [20]). What is more, such entrainment of linguistic representations is said to form the basis for mutual understanding [21].

(a) Coordination of spatial and temporal neural activity between speakers and listeners

Corresponding to the assumption of a shared representational format for comprehending and producing language, on a neural level, neural circuitry underlying language comprehension and production have been shown to overlap (e.g. [22–24]). This provides the basis for hypothesizing that neural states of communicating individuals should become more similar to each other as they activate similar representations while speaking and listening. This was first demonstrated by a seminal functional magnetic resonance imaging (fMRI) study comparing the brain activity of a speaker telling a story to the brain activity of several listeners hearing the story [25]. By relating the two patterns of neural activity to each other the authors were able to demonstrate that speaker's and listeners' brain activity coordinated both temporally and spatially. Temporally, listeners' brain activity mainly followed speakers' brain activity with a delay of several seconds as the signal passed from speaker to listeners. A delay in neural coordination has since been also found in other dual-brain studies and has been related to different levels in the processing hierarchy (e.g. [26,27]) and to communicative success (e.g. [28]). In Stephens and colleagues' study [25] listeners' brain activity preceded speaker's brain activity in some cases, which the authors interpreted as anticipatory activation (for more recent findings on the role of prediction in speaker–listener coordination, see [29]).

Spatially, interpersonal neural coordination was observed in corresponding brain areas in speakers and listeners, providing support for the representational similarity between these two processes. Interpersonal neural coordination is associated to communicative success: the degree of spatial coordination between speaker and listener correlated with listeners' story comprehension, and, when the story was told in a language the listeners could not understand, no significant coupling was observed.

A coordination between speakers and listeners may not exclusively rely on activating identical brain areas and processes: in a similar storytelling setting, a dual-brain

electroencephalography (EEG) study reports coordination between speaker's and listeners' electrophysiological activity involving different locations across the skull ([30]; for comparable findings of coordination between different brain areas in nonverbal interaction, see [31]). This study's design explicitly excluded the possibility that neural coordination between speakers and listeners results merely from processing a shared perceptual environment. Instead neural coordination corresponds to processing the content of the communicated linguistic information: the recorded narrations of two speakers were superimposed on each other and listeners were instructed to attend to either one or the other speaker. Listeners' EEG coordinated with the attended speaker's EEG significantly more than with the unattended speaker's EEG.

This is in line with more recent contributions, which have argued that mutual understanding is more than a one-directional flow of information, encoding a message on the one side and decoding the message on the other side. Instead, mutual understanding is actively generated and involves dynamic adaptation and the deliberate building of a shared conceptual space between speakers and listeners [32–35]. That mutual understanding goes beyond producing and receiving speech signals (and activating corresponding representations) is also apparent by dual-brain studies indicating a speaker–listener coordination in brain areas that extend beyond the core language network and engages additional cognitive processes, such as working memory [36], shared attention [37] and social cognition [25,30,31,38].

Studies that relate neural activity across multiple individuals have pioneered methodological advances that allow investigating and quantifying neural coordination between two or more brains (e.g. [39–41]). Multi-brain studies can offer insights that go beyond the individual mind and brain, and investigate the dynamic coordination between two or more communicating individuals. Such an approach can further theory-building and theory-testing, and advance our understanding of language in social interaction (see [42] for a recent review of this research area).

(b) Coordination of behaviour and cognition during turn-taking

Another feature of turn-taking that has sparked a rapidly growing body of experimental research, is the speed by which speakers and listeners alternate. A starting point for these investigations has been the observation that, across cultures, the average inter-turn interval, the gap in between one speaker's speaking turn and the next speaker's turn, is around 200 ms [43,44]. This is in stark contrast with psycholinguistic laboratory studies on language production in which the production of a word (let alone a sentence) takes already 600 ms (e.g. [45,46]). This puzzle has been addressed by assuming a cognitive architecture that can comprehend and produce language in parallel and relies on predictive mechanisms that allow the next speaker to prepare their speaking turn once the gist of the incoming turn can be predicted and to launch it once the current turn is completed (e.g. [47–52]).

A seminal study of turn-taking manipulated how early an interlocutor can predict how their conversational partner's speaking turn will end and hence, how early the planning of a response can begin [49]. Resorting to recordings of electrophysiological activity in the interlocutor's brain the

researchers were able to demonstrate that speech planning begins already during the conversational partner's speaking turn, and begins earlier when the gist of the partner's speaking becomes predictable earlier in the sentence. Comparable conclusions have been reached by Corps *et al.* [51]. Noteworthy is also a recent study by Bögels [50] which replicated the finding that conversational partners can and do start planning their turn already about one-third into the partner's utterances in a more natural setting with a mix of pre-scripted and spontaneous turns.

While there seems to be agreement in the field that speech planning can and does occur in parallel to comprehending the partner's speech, it is not clear to what degree the two processes may conflict with each other. If a person begins to plan their response while their partner is still speaking, this may reduce the depth with which the partner's utterances can be processed. First investigations have addressed the costs of early planning on processing load (e.g. [53,54]). The theoretical basis of these investigations is that preparing speaking while simultaneously processing incoming speech taps into the same mental representations, limiting the resources that can be dedicated to each. Yet a recent neurosurgical study using intracranial electrocortical recordings comes to the conclusion that speech planning is functionally and anatomically distinct from articulating speech or listening [55]. This study temporally isolated the initiation of speech planning processes by varying the point at which critical information is presented in the partner's turn (cf. [49]), and then compared the observed neural activity to neural activity observed during episodes of listening and speaking. The identified signatures of speech-selective planning were then verified in unconstrained natural dialogue. If indeed, as this study suggests, separate modules are dedicated to speech planning and speech processing, this could mean that the two processes can run in parallel without impeding each other.

A recent analysis of corpus data of naturalistic spoken conversations challenges the proposal that preparing speech production during comprehension is pervasive in everyday conversation [56]. An analysis of the length of speech segments suggests that speaking turns may often not be long enough to allow their partners to complete, or even initiate, the planning of their response. According to the authors, further strategies may be in place to assure timely turn-taking. For example, interlocutors may not respond to the immediately preceding speaking turn, but to a topic introduced earlier in the conversation. Essentially, interlocutors may each be contributing to the conversation but may not be directly referring to each other's contributions. This proposal fits with the proposal that conversational partners may strive for a level of understanding that is good enough for the current purpose [57,58], and which may sometimes be built on an illusion of understanding [59].

Lastly, another open question arising from this research field is which cues speakers rely on for predicting their partner's turn end so that they can launch their response. This response needs to be not only appropriate in content [60] but also appropriate in time (e.g. [61–66]) as the time interval that elapses between two speaking turns carries pragmatic information (e.g. making certain replies more likely than others, see [67]). Different linguistic and paralinguistic cues have been demonstrated to mark upcoming turn ends, for example, syntactic and lexical markers [68] or prosodic

properties of speech [62], and may be used to facilitate turn transitions. Lastly, it has been proposed that the inherent temporal structure of conversation itself may facilitate the timing of speaking turns [66,69]. For example, a conversational partner may rhythmically entrain on the current speaker's syllable rate. Applying a dynamical systems perspective, the temporal coordination observed during turn-taking may thus be described by self-organizing dynamics between coupled oscillators. Similar proposals have also been made for non-verbal types of social interaction, such as finger tapping [70], or body sway [71].

While theory-building in this research field continues, the question of how speaking and listening become so tightly interweaved in social interaction remains valid and is an important puzzle piece to theories of language production in social interaction. Next we will turn to another demand social interaction places on language production, namely going beyond actions of individuals and merging these into a joint action.

4. Representing the own and the partner's actions

Language production, when used for social interaction, is not simply an individual process. Speaking becomes a coordinated, joint action between two or more socially interacting individuals (e.g. [1,3,18,72–74]).

Over the last decades, this perspective has gained from cross-fertilization from research on the perceptual, motor, and cognitive basis of joint actions (for overview, see [75]). Joint action has been defined as an action for which two or more individuals coordinate their individual behaviour in time and space to achieve a shared goal [76]. One prominent theoretical proposal has been that coordination between multiple individuals relies on an individual's capacity to represent the joint task, to predict the outcome of the own and the partner's contributions, and to integrate these [77]. Empirical support for this comes from studies on action planning in a joint spatial compatibility task, demonstrating that co-acting individuals represent not only their own actions but also their partner's actions (e.g. [78]; for recent meta-analysis, [79]).

These theoretical assumptions have recently been applied also to joint actions involving speaking. A seminal study by Baus *et al.* [80] recorded electrophysiological activity (EEG) during a joint picture naming task in which two task partners took turns naming pictures. Using electrophysiological signatures of lexical frequency, the authors were able to demonstrate that speakers engage in lexical processing not only when it is their turn to name a picture but also when it is their partner's turn to name the picture. The authors propose that representing the partner's speaking enables interlocutors to predict their partner's verbal behaviour. This is in line with recent mechanistic accounts of dialogue which propose that predicting the partner's verbal behaviour engages the own speech production system and serves to facilitate language processing and the coordination of utterances across speakers both in their timing and content (e.g. [18,81,82]).

An ongoing debate in the field of joint action research is the nature of these shared task representations. While some propose that the partner's specific action is simulated and represented on a level of detail similar to one's own

action [83,84], others have proposed that represented is not the partner's action itself but merely the fact that the partner acts [85–89].

Recent speech production studies investigating language production in a joint action framing can be applied to this debate; albeit providing mixed evidence. First evidence in favour of the proposal that the partner's speaking is represented as detailed as own speaking comes from the study by Baus *et al.* discussed above: since the neurophysiological activity elicited by the partner's naming were sensitive to lexical frequency this suggests that at least this aspect of the partner's action was represented. Further evidence comes from a study by us [90] in which we built upon a well-known semantic interference effect characterized by slower picture naming when successively naming pictures of the same semantic category [91]. Placed in a joint picture naming setting in which two task partners take turns naming pictures, this study was able to demonstrate that semantic interference can be elicited not only by own prior naming of semantically related pictures but also by the partner's naming of semantically related pictures. Importantly, partner-elicited semantic interference was not based on hearing the partner name the picture: the same pattern of results emerged when the partner's naming was masked through noise-canceling headphones (Kuhlen & Abdel Rahman, experiment 3), and when task partners were naming pictures located in different rooms (Kuhlen & Abdel Rahman, experiment 2). This pattern of results supports proposals that the partner's utterances are simulated based on the beliefs about the partner's action and, crucially that this simulation is at the level of seeking lexical access on behalf of the partner. Consequently, the amount of interference experienced increases when subsequently naming semantically related pictures. Converging evidence for this finding comes from a different team of researchers who applied the same paradigm also in a joint picture naming setting [92] and from a mega-analysis, combining several studies within this setting [93]. These studies align with accounts of joint action which suggest that the partner's actions are co-represented and suggests that the partner's speaking can have profound and lasting effects on own speech production.

Other studies, however, find little evidence for the claim that the partner's utterances are co-represented—at least not to the level of seeking lexical access on behalf of the partner. For example, a study by Gambi *et al.* [94] found that speech production was affected by the knowledge that the speakers' task partner was concurrently also naming a picture. But speakers were not affected by whether the picture was related or unrelated to the pictures the partner had to name [94]. Similar conclusions have been reached by other studies employing variants of joint picture naming tasks [95,96].

Whether a partner's utterances are represented, and to what level of detail, may depend on the nature of the social setting, the role the task partner plays in the interaction, and the saliency of their actions. For example, in a follow-up study on the cumulative semantic interference effect in a joint picture naming setting we were not able to find partner-elicited interference [97]. In this study we recorded electrophysiological activity of the speakers, which required a large number of trials and, to avoid the confound of hearing the partner name the picture, required that the two task partners completed the joint picture naming task sitting in two separate rooms. This may have weakened the effect: a mega-analysis

accumulating six experiments on cumulative partner-elicited semantic interference suggests that partner-effects may be particularly pronounced after just having interacted (in physical presence) with the partner [93]. How a partner's actions are represented may also depend on the identity of the task partner: studies on human–robot interaction indicate that a robot's verbal actions may be represented on the conceptual, but not on the lexical level ([98]; more on conceptual alignment with robot partners [99]). Lastly, literature from research on other types of joint action suggests that the goals of the individual co-actors, whether they complement or compete with another, may influence partner co-representations [100,–102]. Future studies will need to further elucidate how these factors affect the degree to which the partner's actions are represented.

The body of research reviewed here has scaled up controlled studies of language production, specifically on the process of lexical selection, to joint action settings in which two speakers take turns speaking. Yet the employed picture naming tasks provide, at best, only very minimal communicative purpose for the speaker's utterances. Speaking with the goal to communicate is arguably an important aspect of language use in social interaction. What is more, conversational speech typically involves more than producing isolated words. Instead these are most often embedded in a larger discourse context. These aspects may shape the processes by which individuals produce language. This is indeed suggested by a recent joint picture naming study in which we embedded a picture–word interference task in a conversational turn sequence [97]. In this study, two participants played a simple game that involved naming and matching pictures displayed on their playing cards: in a given trial, the first player requests the information which cards should be placed on top of each other (e.g. 'which card comes on apple?') thereby producing the distractor word (apple). The second player then responded by naming the appropriate card thereby naming the target picture that was either semantically related or unrelated to the distractor (e.g. 'pear' versus 'car'). Hence, the experimental procedures essentially resembled a classic picture–word interference task (compare e.g. [103–105]), but placed in a social and communicatively meaningful exchange. The large body of literature on picture–word interference in single-speaker settings would predict that speakers should experience semantic interference, that is slower naming latencies, when hearing a semantically related distractor word immediately prior to naming the target picture. Yet, in this communicative setting, no interference was observed. What is more, when the communicative game was slightly altered to increase participants' focus on the conceptual relationship between target and distractor (e.g. 'what matches apple?'), the partner's utterance, when presented 650 ms before the onset of the target picture, even had the potential to prime and facilitate subsequent speaking. We propose that, in communicative settings, the partner's utterances are considered in relationship to own utterances and can thus facilitate own speaking. In the framework of picture–word interference, this means a shift towards processing the distractor words in its conceptual relationship to the target word (instead of focusing on its lexical representation), thus moving from lexical interference to semantic priming.

Along similar lines, other experiments embedding picture–word interference in a larger discourse context have also reported decreased semantic interference ([106,107], experiment 1). Yet, the mechanism behind this effect is not quite

clear. While we have argued that the decrease of semantic interference results from a greater emphasis on conceptual processing and enhances semantic priming [97], Shao & Rommers [107] suggest the preceding discourse context constrains semantic processing by narrowing down the number of possible lexical candidates (but see [108,109], discussed below, for empirical evidence against this degree of flexibility). Tufft & Richardson [106] propose a non-linguistic origin for decreased semantic interference, resulting from the possibility of offloading responsibility for processing the distractor in a dual-task setting. While the precise mechanism of this effect requires further investigation, these studies suggest that the communicative nature of the social interaction can have an impact on how the partner's utterances are integrated with, and shape, own language production.

5. Adapting speaking to the conversational partner

When speaking in a social setting, utterances are adapted to the specific conversational partner. Such adaptations have been called audience or recipient design [110–112] and can shape all levels of linguistic representations: for example, the choice of a referential expression depending on the partner's perspective (e.g. [113–120]). The choice of syntactic constructions depending on information being new or already known to the addressee (e.g. [121,122]), and the clarity with which this information is articulated [122,123]. These types of flexible adaptations in language use have been suggested to be supported by the human ability to take another person's perspective and to infer their mental states [1]; an ability typically referred to as mentalizing or Theory of Mind. The ability to mentalize may be an important additional process engaged when language is produced in social interaction. In addition to engaging social cognition, adaptations to the conversational partner (and the conversational context more generally) may trigger adaptations in semantic processing. In the following we will discuss empirical evidence for each of these aspects underlying partner-adapted language production.

(a) Mentalizing supports partner-adapted speech production

Numerous studies on speech comprehension have pointed towards the need to engage in pragmatic inferencing in order to derive from linguistic meaning intended speaker meaning (for review [124]). These inferences are typically associated with social cognition, and a person's ability to read another person's (communicative) intentions and mental states, also called mentalizing or Theory of Mind.

Also when speaking, mentalizing is routinely engaged. For example, an fMRI study investigated neural activity of speakers while engaged in a communicative game with a task partner located outside of the scanner [125]. Specifically, the study contrasted a communicative condition, in which speakers referred to objects whose identity their task partner had to guess, with a non-communicative condition, in which the addressee knew in advance which object was being referred to (thereby eliminating any need to communicate this information). Areas that social neuroscience has associated with mentalizing, most notably the dorsomedial prefrontal cortex, differentiated these conditions. The authors concluded that

mentalizing is a crucial process when communicating to a conversational partner.

Similar conclusions are reached by another fMRI study in which speakers were asked to give simple spatial instructions either to a task partner located outside of the scanner (communicative condition), or for the purpose of testing the microphone (non-communicative condition; [126]). Neural activity was measured prior to speaking at a time point when speakers knew that they would be speaking with or without communicative intentionality ('who': partner versus microphone test), but they did not know yet which instructions they should give ('what'). Multivariate pattern analysis decoding regional activation patterns associated with a particular task set [127–129] revealed that the communicative intentionality of the setting is encoded in the ventral-medial prefrontal cortex. This brain area has been identified in lesion studies to be responsible for tailoring a communicative message to characteristics of the task partner [130]. This study supports the conclusion that mentalizing is routinely engaged during speaking in social interaction and that mentalizing becomes engaged already during preparatory processes of speaking.

The assumption that pragmatic processes shape speaking already in preparatory stages of speech production is further supported by a recent EEG study in which participants were asked to speak with the intention of producing two different types of speech acts—either for the purpose of naming the displayed object or with the intention to request the object [131]. Electrophysiological signatures, most notably a readiness potential that the authors label pragmatic prediction potential, distinguishes these two speech acts already 600 ms prior to speaking. Based on a source analysis of the observed predictive potential, which indicates the involvement of the motor cortex, the authors suggest that preparing a request engages predictions about specific action-related consequences. This study supports accounts that a speaker's communicative intentions, and possibly the anticipation of the addressee's reaction, inform preparatory processes of language production.

The above reviewed studies indicate that pragmatic processes, most notably the ability to infer and anticipate another person's intention and mental states, become engaged when speakers speak in the context of a communicative, social exchange. A recent fMRI study ties activation of the mentalizing network to the ability to adapt speaking to a conversational partner's informational needs [132]. In this study participants, located inside the scanner, gave their task partner, located outside the scanner, instructions to select objects placed on a grid, a classic referential communication game for investigating audience design (see [133–135]). In a communicative condition speakers' utterances served the purpose to instruct the task partner which object to select. This condition was compared to a noncommunicative condition, in which speakers produced the referential expressions without having a task partner. Neural activity was recorded during speech planning, which was experimentally separated from speech production by requesting participants to press a button prior to speaking. Crucially, in some trials the grid contained objects that competed with the target object (a small glass and a medium-sized glass), and these were either visually accessible, or occluded from their partner's perspective on the grid. This required speakers to specify their referring expression depending on their partner's perspective. The

pattern of results suggests the involvement of two core structures of the mentalizing network, the medial prefrontal cortex (mPFC) and the temporal parietal junction (TPJ), when speakers produced speech for the purpose to communicate and when their utterances required an adaptation to the partner's perspective. These two brain areas may serve a differential role: activity in the mPFC was observed under communicative (versus noncommunicative) conditions in which speakers had to adjust their utterances to their task partner's perspective. Activity in the bilateral TPJ was exclusively observed within communicative trials and marked whether the speaker's and their addressee's perspective differed. Together these findings support theories that propose the mPFC distinguishes communicative from noncommunicative actions (e.g. [125,126]) and further suggests that the TPJ supports flexibly adaptations to the specific needs of the addressee.

One disputed question in the field has been how essential mentalizing is to speaking in social settings. While the just reviewed work suggests that mentalizing is routinely engaged already during early moments of language production, other scholars have proposed that speaking is initially planned based on the speaker's own perspective and mentalizing is employed only later in processing in case a repair is needed (e.g. [136–140]). Neuroscientific support for this claim comes from a magnetoencephalography study on language comprehension in which addressees processed referential expressions that had either been established as common ground with the speaker in a prior interaction, or that violated this common ground [141]. Mentalizing activity was only observed when speakers violated the common ground established with their addressee. The authors conclude that mentalizing is engaged as a reaction to pragmatic violations, but not for making partner-specific predictions about the partner's upcoming utterances.

While the reviewed studies on speech production suggest that mentalizing is an essential component when speakers speak with communicative intentionality and allows speakers to adapt their speaking to the conversational context, the connection between brain areas associated with speaking and those associated with mentalizing are not yet well understood. On one hand, the literature supports the view that mentalizing and linguistic processing are separate faculties: for example, speakers with aphasia show difficulties speaking while their mentalizing skills are typically intact (e.g. [142–146]); and, vice-versa, speakers with autism spectrum disorder show deficits in mentalizing while typically their linguistic abilities are intact (e.g. [147–152]). On the other hand, the here reviewed studies suggest a close integration of the ability to speak and the ability to mentalize. This implies a connection between areas supporting language production and those supporting mentalizing. Indeed, Deen *et al.* [153] report a small amount of overlap between language and mentalizing activations in the left superior temporal cortex. Moreover, a recent study involving three fMRI experiments reports synchronized activity between language and mentalizing networks at rest and during story comprehension [154], suggesting a functional integration of these two abilities. More work is needed, though, to understand the nature of this integration and how it plays out over the course of producing a communicative utterance. Employing more time-sensitive neuroscientific methods, such as EEG, could allow future research to address these questions.

(b) Knowledge about the partner shapes semantic processing

Partner-adapted speaking may not only involve thinking about the communication partner's mental states and informational needs, it may also carry consequences for semantic processing. Along these lines, research on language comprehension has demonstrated that semantic processing is shaped by knowledge about the conversational partner: for example, listeners use cues available in the speaker's accents to disambiguate homonyms [155], and the speakers' voice or face can serve as a cue to the speaker's identity and inform predictions about upcoming words [156,157]. Also, knowledge about the partner's profession can shape word processing such that atypical category exemplars become more accessible [158]. These studies point towards a flexibility in semantic processing that enable adapting language comprehension to the conversational partner.

Similar flexibility may be observed in language production. In support of this proposal it has been demonstrated that semantic relationships can be constructed on the fly given a certain semantic context and that these *ad-hoc* semantic relationships affect speech production [159]. While objects like 'stool', 'bucket', 'knife' and 'river' are typically perceived as semantically rather unrelated, they are perceived as related when presented together with an overarching theme, in this case, a fishing trip. Once this theme is available to speakers the generated semantic relationship between these items leads to interference during lexical selection. The same principle may be at work in a conversational context. First evidence for this comes from a study demonstrating that thematic context can be elicited by information provided by a task speaker [160]. In this study formerly unrelated objects were associated to a common theme through a short narrative provided by the task partner (e.g. a narrative about going on a fishing trip). Similar to Abdel Rahman & Melinger's study [159], the thematic context generated by the partner elicited, under certain circumstances, interference in a subsequent picture naming task.

Yet other studies have pointed to limited flexibility of speakers to take the social or pragmatic context into account. In a joint picture naming study [161], speakers took turns naming pictures. In some trials, the preferred basic level term for a given picture became pragmatically infelicitous because it failed to distinguish between two objects of the same category (e.g. naming the picture of a shark 'fish' in the presence of another fish). Yet, when words phonologically related to this basic level term were presented as distractors in a picture-word interference paradigm, they interfered with speech production. This indicates that lexical entries for the basic level term were covertly activated during speech production even though they were pragmatically infelicitous, suggesting that the semantic system shows little flexibility in taking pragmatic context into account (for comparable findings, see [108]).

Future studies will need to pinpoint the degree of flexibility, and its limits, for pragmatic and social context to constrain semantic processing during language production. This will inform models of language production and help distinguish accounts that view word retrieval as a largely static process (e.g. [162–170]) from accounts that allow for more context-dependent flexibility (e.g. [171,172]).

6. How do we get to a framework for language production in social interaction

Speaking is a fundamentally social activity: when we speak, we typically speak in the context of a social interaction. In order to do so successfully, our cognitive system does not only need to solve the challenge of producing fluent speech; it must also meet the demands social interaction places on the interacting individuals. In this paper we have focused on three main demands: (i) the need to integrate speaking with listening, (ii) the need to build from single actions a joint action, and (iii) the need to flexibly adapt to the conversational partner.

To meet these demands, additional processes are engaged that go beyond the core faculty of speaking. To allow for timely turn-taking, speakers must plan their speech and, in parallel, process the incoming speech from their partner's turn. Furthermore, speakers plan, represent and execute not only their own verbal actions, they also predict and represent their partner's actions in order to coordinate, in time and content, their own and their partner's behaviour. This requires taking into account the partner's unique perspective, knowledge background and informational needs, involving processes of social cognition and mentalizing. Speaking in social interaction therefore includes a series of additional cognitive processes. This argument is in line with other recent calls for extending the traditional language network [124,173,174].

What is more, the processes of language production are shaped by social interaction. The verbal actions of the partner can have a profound and lasting effect on own language production. For example, the utterances of the partner contribute to the current semantic context and can as such interfere with a speaker's lexical access [90,92], or facilitate this access [97,106,107]. Moreover, a given semantic context can be shaped by knowledge about or information shared with the conversational partner [160]. Language comprehension may be facilitated when combined with episodes of speaking (e.g. [81,175]) and may be affected by the knowledge state of the conversational partner (e.g. [176]).

In order to develop a theoretical account of language production in social interaction it is necessary to integrate our understanding of language production with our understanding of how people understand others' intentions and mental state and how multiple actors coordinate their behaviour and mental states in social interactions more generally. Indeed, in recent years these formerly separate research areas have grown together as cognitive psychologists are becoming increasingly interested in social settings; and social psychologists are becoming increasingly interested in cognitive accounts for social behaviour [177].

Advancing this field requires methodological innovations that allow experimental investigations of language use in social settings. This may require innovating or adapting existing experimental procedures and designs so that two or more individuals can interact in controlled laboratory settings. Some questions may be addressed best by recording not only one individual's behaviour, cognition or neural states but also of those of the interaction partner. This shadows recent calls for a paradigmatic shift in social neuroscience to a two-person approach (e.g. [40,178,179]).

Methodological advances also require more interactive settings that allow research participants to actively engage in social interactions instead of just passively observing them [180–182]. This development is also met by recent

calls for a shift in the study of social cognition towards more ecologically valid settings in dynamic, multimodal, context-embedded and interactive environments [183]. Last but not least, when bringing social interaction into the laboratory, the research participants' behaviour is influenced not only by the presented stimuli, but also by the interaction partner. Researchers are therefore cautioned to consider, just as carefully as they consider the selection of their stimuli, how an interaction partner is introduced into the experimental setting [73].

7. Conclusion

Investigating language production in socially isolated laboratory settings has laid the ground works for understanding the cognitive and neural underpinnings of speaking. Language production in social interaction is not fundamentally different

from language production in isolation. Yet social interaction places important constraints on language production that are met by engaging additional processes and shape the processes underlying language production. It is time to scale up our understanding of language production to settings in which two or more speakers communicate in social interaction.

Data accessibility. This article has no additional data.

Authors' contributions. A.K.K.: conceptualization, writing—original draft; R.A.R.: writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This research was supported by grants KU3236/3 and AB277/11 from the German Research Council and the Career Development Award 420_CDA_7 awarded to the first author by the Berlin University Alliance.

References

- Brennan SE, Galati A, Kuhlen AK. 2010 Two minds, one dialog. In *Psychology of learning and motivation*, vol. 53 (ed. BH Ross), pp. 301–344. Amsterdam, The Netherlands: Elsevier.
- Tanenhaus MK, Brown-Schmidt S. 2008 Language processing in the natural world. *Phil. Trans. R. Soc. B* **363**, 1105–1122. (doi:10.1098/rstb.2007.2162)
- Pickering MJ, Garrod S. 2004 Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* **27**, 169–190. (doi:10.1017/S0140525X04000056)
- Levelt WJM. 1989 *Speaking: from intention to articulation*. Cambridge, MA: The MIT Press.
- Meyer AS, Roelofs A, Brehm L. 2019 Thirty years of speaking: an introduction to the Special Issue. *Lang. Cogn. Neurosci.* **34**, 1073–1084. (doi:10.1080/23273798.2019.1652763)
- Holler J, Levinson SC. 2019 Multimodal language processing in human communication. *Trends Cogn. Sci.* **23**, 639–652. (doi:10.1016/j.tics.2019.05.006)
- Goodwin C. 1981 *Conversational organization: interaction between speakers and hearers*. New York, NY: Academic Press.
- Jefferson G. 1973 A case of precision timing in ordinary conversation: overlapped tag-positioned address terms in closing sequences. *Semiotica* **9**, 47–96. (doi:10.1515/semi.1973.9.1.47)
- Sacks H, Schegloff EA, Jefferson G. 1974 A simplest systematics for the organization of turn-taking for conversation. *Language* **50**, 696. (doi:10.2307/412243)
- Levinson SC. 2016 Turn-taking in human communication – origins and implications for language processing. *Trends Cogn. Sci.* **20**, 6–14. (doi:10.1016/j.tics.2015.10.010)
- Prinz W. 1990 A common coding approach to perception and action. In *Relationships between perception and action* (eds O Neumann, W Prinz), pp. 167–201. Berlin, Germany: Springer.
- Gallese V, Fadiga L, Fogassi L, Rizzolatti G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609. (doi:10.1093/brain/119.2.593)
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L. 1996 Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141. (doi:10.1016/0926-6410(95)00038-0)
- Rizzolatti G, Fogassi L, Gallese V. 2001 Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670. (doi:10.1038/35090060)
- Dell G, Chang F. 2014 The P-Chain: relating sentence production and its disorders to comprehension and acquisition. *Phil. Trans. R. Soc. B* **369**, 20 120 394. (doi:10.1098/rstb.2012.0394)
- Huettig F. 2015 Four central questions about prediction in language processing. *Brain Res.* **1626**, 118–135. (doi:10.1016/j.brainres.2015.02.014)
- Pickering MJ, Garrod S. 2007 Do people use language production to make predictions during comprehension? *Trends Cogn. Sci.* **11**, 105–110. (doi:10.1016/j.tics.2006.12.002)
- Pickering MJ, Garrod S. 2013 An integrated theory of language production and comprehension. *Behav. Brain Sci.* **36**, 329–347. (doi:10.1017/S0140525X12001495)
- Pardo JS. 2006 On phonetic convergence during conversational interaction. *J. Acoust. Soc. Am.* **119**, 2382–2393. (doi:10.1121/1.2178720)
- Mol L, Krahmer E, Maes A, Swerts M. 2012 Adaptation in gesture: converging hands or converging minds? *J. Mem. Lang.* **66**, 249–264. (doi:10.1016/j.jml.2011.07.004)
- Pickering MJ, Garrod S. 2006 Alignment as the basis for successful communication. *Res. Lang. Comput.* **4**, 203–228. (doi:10.1007/s11168-006-9004-0)
- Buchsbaum BR, Hickok G, Humphries C. 2001 Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cogn. Sci.* **25**, 663–678. (doi:10.1207/s15516709cog2505_2)
- Menenti L, Gierhan SME, Segaert K, Hagoort P. 2011 Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychol. Sci.* **22**, 1173–1182. (doi:10.1177/0956797611418347)
- Silbert LJ, Honey CJ, Simony E, Poeppel D, Hasson U. 2014 Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proc. Natl Acad. Sci. USA* **111**, 43. (doi:10.1073/pnas.1323812111)
- Stephens GJ, Silbert LJ, Hasson U. 2010 Speaker–listener neural coupling underlies successful communication. *Proc. Natl Acad. Sci. USA* **107**, 14 425–14 430. (doi:10.1073/pnas.1008662107)
- Chang CHC, Nastase SA, Hasson U. In press. Information flow across the cortical timescales hierarchy during narrative construction. *Neuroscience*. (doi:10.1101/2021.12.01.470825)
- Kuhlen AK, Allefeld, Haynes JD 2012 Content-specific coordination of listeners' to speakers' EEG during communication. *Front. Hum. Neurosci.* **6**, 266. (doi:10.3389/fnhum.2012.00266)
- Davidesco I, Laurent E, Valk H, West T, Dikker S, Milne C, Poeppel D. In press. Brain-to-brain synchrony between students and teachers predicts learning outcomes. *Neuroscience*. (doi:10.1101/644047)
- Dikker S, Silbert LJ, Hasson U, Zevin JD. 2014 On the same wavelength: predictable language enhances speaker-listener brain-to-brain synchrony in posterior superior temporal gyrus. *J. Neurosci.* **34**, 6267–6272. (doi:10.1523/JNEUROSCI.3796-13.2014)
- Kuhlen AK, Allefeld C, Haynes J-D. 2012 Content-specific coordination of listeners' to speakers' EEG during communication. *Front. Hum. Neurosci.* **6**, 266. (doi:10.3389/fnhum.2012.00266)
- Dumas G, Nadel J, Soussignan R, Martinerie J, Garnero L. 2010 Inter-brain synchronization during

- social interaction. *PLoS ONE* **5**, e12166. (doi:10.1371/journal.pone.0012166)
32. Brown-Schmidt S, Duff MC. 2016 Memory and common ground processes in language use. *Top. Cogn. Sci.* **8**, 722–736. (doi:10.1111/tops.12224)
 33. Fusaroli R, Tylén K. 2016 Investigating conversational dynamics: interactive alignment, interpersonal synergy, and collective task performance. *Cogn. Sci.* **40**, 145–171. (doi:10.1111/cogs.12251)
 34. Schober MF. 2004 Just how aligned are interlocutors' representations? *Behav. Brain Sci.* **27**, 209–210. (doi:10.1017/S0140525X04420056)
 35. Stolk A, Verhagen L, Toni I. 2016 Conceptual alignment: how brains achieve mutual understanding. *Trends Cogn. Sci.* **20**, 180–191. (doi:10.1016/j.tics.2015.11.007)
 36. Ahn S, Cho H, Kwon M, Kim K, Kwon H, Kim BS, Chang WS, Chang JW, Jun SC. 2018 Interbrain phase synchronization during turn-taking verbal interaction: a hyperscanning study using simultaneous EEG/MEG: synchronization during turn-taking verbal interaction. *Hum. Brain Mapp.* **39**, 171–188. (doi:10.1002/hbm.23834)
 37. Gvirts HZ, Perlmutter R. 2020 What guides us to neurally and behaviorally align with anyone specific? A neurobiological model based on fNIRS hyperscanning studies. *Neuroscientist* **26**, 108–116. (doi:10.1177/1073858419861912)
 38. Hirsch J, Adam Noah J, Zhang X, Dravida S, Ono Y. 2018 A cross-brain neural mechanism for human-to-human verbal communication. *Soc. Cogn. Affect. Neuroscience* **13**, 907–920. (doi:10.1093/scan/nsy070)
 39. Hamilton AFdC. 2021 Hyperscanning: beyond the hype. *Neuron* **109**, 404–407. (doi:10.1016/j.neuron.2020.11.008)
 40. Konvalinka I, Roepstorff A. 2012 The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Front. Hum. Neurosci.* **6**, 215. (doi:10.3389/fnhum.2012.00215)
 41. Kuhlen AK, Allefeld C, Anders S, Haynes J-D. 2015 Towards a multi-brain perspective on communication in dialogue. In *Cognitive neuroscience of natural language use* (eds RM Willems), pp. 182–200, 1st ed. Cambridge, UK: Cambridge University Press.
 42. Kelsen BA, Sumich A, Kasabov N, Liang SHY, Wang GY. 2022 What has social neuroscience learned from hyperscanning studies of spoken communication? A systematic review. *Neurosci. Biobehav. Rev.* **132**, 1249–1262. (doi:10.1016/j.neubiorev.2020.09.008)
 43. Heldner M, Edlund J. 2010 Pauses, gaps and overlaps in conversations. *J. Phonetics* **38**, 555–568. (doi:10.1016/j.wocn.2010.08.002)
 44. Stivers T *et al.* 2009 Universals and cultural variation in turn-taking in conversation. *Proc. Natl Acad. Sci. USA* **106**, 10 587–10 592. (doi:10.1073/pnas.0903616106)
 45. Ferreira F. 1991 Effects of length and syntactic complexity on initiation times for prepared utterances. *J. Mem. Lang.* **30**, 210–233. (doi:10.1016/0749-596X(91)90004-4)
 46. Indefrey P, Levelt WJM. 2004 The spatial and temporal signatures of word production components. *Cognition* **92**, 101–144. (doi:10.1016/j.cognition.2002.06.001)
 47. Barthel M, Levinson SC. 2020 Next speakers plan word forms in overlap with the incoming turn: evidence from gaze-contingent switch task performance. *Lang. Cogn. Neurosci.* **35**, 1183–1202. (doi:10.1080/23273798.2020.1716030)
 48. Barthel M, Sauppe S, Levinson SC, Meyer AS. 2016 The timing of utterance planning in task-oriented dialogue: evidence from a novel list-completion paradigm. *Front. Psychol.* **7**, 1858. (doi:10.3389/fpsyg.2016.01858)
 49. Bögels S, Magyari L, Levinson SC. 2015 Neural signatures of response planning occur midway through an incoming question in conversation. *Sci. Rep.* **5**, 12881. (doi:10.1038/srep12881)
 50. Bögels S. 2020 Neural correlates of turn-taking in the wild: response planning starts early in free interviews. *Cognition* **203**, 104347. (doi:10.1016/j.cognition.2020.104347)
 51. Corps RE, Crossley A, Gambi C, Pickering MJ. 2018 Early preparation during turn-taking: listeners use content predictions to determine what to say but not when to say it. *Cognition* **175**, 77–95. (doi:10.1016/j.cognition.2018.01.015)
 52. Levinson SC, Torreira F. 2015 Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology* **6**, 1858. (doi:10.3389/fpsyg.2015.00731)
 53. Bartolozzi F, Jongman SR, Meyer AS. 2020 Concurrent speech planning does not eliminate repetition priming from spoken words: evidence from linguistic dual-tasking. *J. Exp. Psychol. Learn. Mem. Cogn.* **47**, 466–480. (doi:10.1037/xlm0000944)
 54. Barthel M, Sauppe S. 2019 Speech planning at turn transitions in dialog is associated with increased processing load. *Cogn. Sci.* **43**, e12768. (doi:10.1111/cogs.12768)
 55. Castellucci GA, Kovach CK, Howard MA, Greenlee JDW, Long MA. 2022 A speech planning network for interactive language use. *Nature* **602**, 117–122. (doi:10.1038/s41586-021-04270-z)
 56. Corps RE, Knudsen B, Meyer AS. 2022 Overrated gaps: inter-speaker gaps provide limited information about the timing of turns in conversation. *Cognition* **223**, 105037. (doi:10.1016/j.cognition.2022.105037)
 57. Clark HH, Wilkes-Gibbs D. 1986 Referring as a collaborative process. *Cognition* **22**, 1–39. (doi:10.1016/0010-0277(86)90010-7)
 58. Ferreira F, Bailey KGD, Ferraro V. 2002 Good-enough representations in language comprehension. *Curr. Dir. Psychol. Sci.* **11**, 11–15. (doi:10.1111/1467-8721.00158)
 59. Lau BKY, Geipel J, Wu Y, Keysar B. 2022 The extreme illusion of understanding. *J. Exp. Psychol.: Gen.* **151**, 2957–2962. (doi:10.1037/xge0001213)
 60. Goreglad Fjaellingsdal T, Schwenke D, Scherbaum S, Kuhlen AK, Bögels S, Meekes J, Bleichner MG. 2020 Expectancy effects in the EEG during joint and spontaneous word-by-word sentence production in German. *Sci. Rep.* **10**, 5460. (doi:10.1038/s41598-020-62155-z)
 61. Brehm L, Meyer AS. 2021 Planning when to say: dissociating cue use in utterance initiation using cross-validation. *J. Exp. Psychol.: Gen.* **150**, 1772–1799. (doi:10.1037/xge0001012)
 62. Bögels S, Torreira F. 2015 Listeners use intonational phrase boundaries to project turn ends in spoken interaction. *J. Phonetics* **52**, 46–57. (doi:10.1016/j.wocn.2015.04.004)
 63. Ruiter J-Pd, Mitterer H, Enfield NJ. 2006 Projecting the end of a speaker's turn: a cognitive cornerstone of conversation. *Language* **82**, 515–535. (doi:10.1353/lan.2006.0130)
 64. Holler J, Kendrick KH, Levinson SC. 2018 Processing language in face-to-face conversation: questions with gestures get faster responses. *Psychon. Bull. Rev.* **25**, 1900–1908. (doi:10.3758/s13423-017-1363-z)
 65. Schaffer D. 1983 The role of intonation as a cue to turn taking in conversation. *J. Phonetics* **11**, 243–257. (doi:10.1016/S0095-4470(19)30825-3)
 66. Wilson M, Wilson TP. 2005 An oscillator model of the timing of turn-taking. *Psychon. Bull. Rev.* **12**, 957–968. (doi:10.3758/BF03206432)
 67. Bögels S, Kendrick KH, Levinson SC. 2015 Never say no ... how the brain interprets the pregnant pause in conversation. *PLoS ONE* **10**, e0145474. (doi:10.1371/journal.pone.0145474)
 68. de Ruiter JP, Mitterer H, Enfield NJ. 2006 Projecting the end of a speaker's turn: a cognitive cornerstone of conversation. *Language* **82**, 515–535.
 69. Pouw W, Holler J. 2022 Timing in conversation is dynamically adjusted turn by turn in dyadic telephone conversations. *Cognition* **222**, 105015. (doi:10.1016/j.cognition.2022.105015)
 70. Konvalinka I, Vuust P, Roepstorff A, Frith CD. 2010 Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Q. J. Exp. Psychol. (Colchester)* **63**, 2220–2230. (doi:10.1080/17470218.2010.497843)
 71. Richardson MJ, Marsh KL, Isenhower RW, Goodman JR, Schmidt RC. 2007 Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Hum. Mov. Sci.* **26**, 867–891. (doi:10.1016/j.humov.2007.07.002)
 72. Brennan SE, Kuhlen AK, Charoy J. 2018 Discourse and dialogue. In *Stevens' handbook of experimental psychology and cognitive neuroscience* (ed. JT Wixted), pp. 1–57. New York, NY: John Wiley & Sons, Inc.
 73. Kuhlen AK, Brennan SE. 2013 Language in dialogue: when confederates might be hazardous to your data. *Psychon. Bull. Rev.* **20**, 54–72. (doi:10.3758/s13423-012-0341-8)
 74. Clark HH. 1996 *Using language*, 1st edn. Cambridge, UK: Cambridge University Press.
 75. Knoblich G, Butterfill S, Sebanz N. 2011 Psychological research on joint action: theory and data. In *The psychology of learning and motivation*:

- advances in research and theory* (ed. BH Ross), pp. 59–101. Amsterdam, The Netherlands: Elsevier Academic Press.
76. Sebanz N, Bekkering H, Knoblich G. 2006 Joint action: bodies and minds moving together. *Trends Cogn. Sci.* **10**, 70–76. (doi:10.1016/j.tics.2005.12.009)
 77. Knoblich G, Butterfill S, Sebanz N. 2011 Psychological research on joint action: theory and data. In *The psychology of learning and motivation*, vol. 54 (ed. BH Ross), pp. 59–101. New York, NY: Academic Press.
 78. Sebanz N, Knoblich G, Prinz W. 2003 Representing others' actions: just like one's own? *Cognition* **88**, B11–B21. (doi:10.1016/S0010-0277(03)00043-X)
 79. Karlinsky A, Lohse KR, Lam MY. 2017 A meta-analysis of the joint Simon effect. In *Proceedings of the 39th annual conference of the cognitive science society* (eds G. Gunzelmann, A. Howes, T. Tenbrink, E.J. Davelaar), pp. 2377–2382. Austin, TX: Cognitive Science Society.
 80. Baus C, Sebanz N, Fuente Vdl, Branzi FM, Martin C, Costa A. 2014 On predicting others' words: electrophysiological evidence of prediction in speech production. *Cognition* **133**, 395–407. (doi:10.1016/j.cognition.2014.07.006)
 81. Baus C, Dubarry A-S, Alario F-X. In press. Electrophysiological modulations of lexical prediction by conversation requirements. *OSF Preprints* (doi:10.31219/osf.io/3cwfz)
 82. Gambi C, Pickering MJ. 2011 A cognitive architecture for the coordination of utterances. *Front. Psychol* **2**, 275. (doi:10.3389/fpsyg.2011.00275)
 83. Atmaca S, Sebanz N, Knoblich G. 2011 The joint flanker effect: sharing tasks with real and imagined co-actors. *Exp. Brain Res.* **211**, 371–385. (doi:10.1007/s00221-011-2709-9)
 84. Sebanz N, Knoblich G, Prinz W. 2005 How two share a task: corepresenting stimulus-response mappings. *J. Exp. Psychol.: Hum. Percept. Perform.* **31**, 1234–1246. (doi:10.1037/0096-1523.31.6.1234)
 85. Dolk T, Hommel B, Colzato LS, Schütz-Bosbach S, Prinz W, Liepelt R. 2011 How "social" is the social Simon effect? *Front Psychol.* **6**, 2–84. (doi:10.3389/fpsyg.2011.00084)
 86. Dolk T, Hommel B, Prinz W, Liepelt R. 2013 The (not so) social Simon effect: a referential coding account. *J. Exp. Psychol.: Hum. Percept. Perform.* **39**, 1248–1260. (doi:10.1037/a0031031)
 87. Philipp AM, Prinz W. 2010 Evidence for a role of the responding agent in the joint compatibility effect. *Q. J. Exp. Psychol. (Colchester)* **63**, 2159–2171. (doi:10.1080/17470211003802426)
 88. Vlainic E, Liepelt R, Colzato LS, Prinz W, Hommel B. 2010 The virtual co-actor: the social Simon effect does not rely on online feedback from the other. *Front. Psychol.* **1**, 208. (doi:10.3389/fpsyg.2010.00208)
 89. Wenke D, Atmaca S, Holländer A, Liepelt R, Baess P, Prinz W. 2011 What is shared in joint action? Issues of co-representation, response conflict, and agent identification. *Rev. Phil. Psychol.* **2**, 147–172. (doi:10.1007/s13164-011-0057-0)
 90. Kuhlen AK, Abdel Rahman R. 2017 Having a task partner affects lexical retrieval: spoken word production in shared task settings. *Cognition* **166**, 94–106. (doi:10.1016/j.cognition.2017.05.024)
 91. Howard D, Nickels L, Coltheart M, Cole-Virtue J. 2006 Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition* **100**, 464–482. (doi:10.1016/j.cognition.2005.02.006)
 92. Hoedemaker RS, Ernst J, Meyer AS, Belke E. 2017 Language production in a shared task: cumulative semantic interference from self- and other-produced context words. *Acta Psychol.* **172**, 55–63. (doi:10.1016/j.actpsy.2016.11.007)
 93. Holtz N, Hauber R, Abdel Rahman R, Kuhlen AK. 2022 Naming together versus naming alone: a mega-analysis of six experiments on joint language production. [Poster presentation] The Biannual Conference of the German Cognitive Science Society, Freiburg, Germany.
 94. Gambi C, Van de Cavey J, Pickering MJ. 2015 Interference in joint picture naming. *J. Exp. Psychol.: Learn. Mem. Cogn.* **41**, 1–21. (doi:10.1037/a0037438)
 95. Brehm L, Taschenberger L, Meyer A. 2019 Mental representations of partner task cause interference in picture naming. *Acta Psychol.* **199**, 102888. (doi:10.1016/j.actpsy.2019.102888)
 96. Hoedemaker RS, Meyer AS. 2019 Planning and coordination of utterances in a joint naming task. *J. Exp. Psychol.: Learn. Mem. Cogn.* **45**, 732–752. (doi:10.1037/xlm000603)
 97. Kuhlen AK, Abdel Rahman R. 2021 Joint language production: an electrophysiological investigation of simulated lexical access on behalf of a task partner. *J. Exp. Psychol.: Learn. Mem. Cogn.* **47**, 1317–1337. (doi:10.1037/xlm0001025)
 98. Wudarczyk OA, Kirtay M, Pischedda D, Hafner VV, Haynes J-D, Kuhlen AK, Abdel Rahman R. 2021 Robots facilitate human language production. *Sci. Rep.* **11**, 16737. (doi:10.1038/s41598-021-95645-9)
 99. Cirillo G, Runnqvist E, Strijkers K, Nguyen N, Baus C. 2022 Conceptual alignment in a joint picture-naming task performed with a social robot. *Cognition* **227**, 105213. (doi:10.1016/j.cognition.2022.105213)
 100. Era V, Aglioti SM, Mancusi C, Candidi M. 2020 Visuo-motor interference with a virtual partner is equally present in cooperative and competitive interactions. *Psychol. Res.* **84**, 810–822. (doi:10.1007/s00426-018-1090-8)
 101. Iani C, Anelli F, Nicoletti R, Rubichi S. 2014 The carry-over effect of competition in task-sharing: evidence from the joint Simon task. *PLoS ONE* **9**, e97991. (doi:10.1371/journal.pone.0097991)
 102. Ruys KI, Aarts H. 2010 When competition merges people's behavior: interdependency activates shared action representations. *J. Exp. Soc. Psychol.* **46**, 1130–1133. (doi:10.1016/j.jesp.2010.05.016)
 103. Damian MF, Martin RC. 1999 Semantic and phonological codes interact in single word production. *J. Exp. Psychol.: Learn. Mem. Cogn.* **25**, 345–361. (doi:10.1037/0278-7393.25.2.345)
 104. Glaser WR, Dünghoff F-J. 1984 The time course of picture-word interference. *J. Exp. Psychol.: Hum. Percept. Perform.* **10**, 640–654. (doi:10.1037/0096-1523.10.5.640)
 105. Schriefers H, Meyer AS, Levelt WJM. 1990 Exploring the time course of lexical access in language production: picture-word interference studies. *J. Mem. Lang.* **29**, 86–102. (doi:10.1016/0749-596X(90)90011-N)
 106. Tufft MRA, Richardson D. 2020 Social offloading: just working together is enough to remove semantic interference. In *Proceedings of the 42nd annual conference of the cognitive science society* (eds S Denison, M Mack, Y Xu, B Armstrong). Austin, TX: Cognitive Science Society Proceedings of the Cognitive Society.
 107. Shao Z, Rommers J. 2020 How a question context aids word production: evidence from the picture-word interference paradigm. *Q. J. Exp. Psychol. (Colchester)* **73**, 165–173. (doi:10.1177/1747021819882911)
 108. Jescheniak JD, Kurtz F, Schriefers H, Günther J, Klaus J, Mädebach A. 2017 Words we do not say—context effects on the phonological activation of lexical alternatives in speech production. *J. Exp. Psychol.: Hum. Percept. Perform.* **43**, 1194–1206. (doi:10.1037/xhp0000352)
 109. Mädebach A, Kurtz F, Schriefers H, Jescheniak JD. 2020 Pragmatic constraints do not prevent the co-activation of alternative names: evidence from sequential naming tasks with one and two speakers. *Lang. Cogn. Neurosci.* **35**, 1073–1088. (doi:10.1080/23273798.2020.1727539)
 110. Bell A. 1984 Language style as audience design. *Lang. Soc.* **13**, 145–204. (doi:10.1017/S004740450001037X)
 111. Clark HH, Carlson TB. 1982 Hearers and speech acts. *Language* **58**, 332. (doi:10.2307/414102)
 112. Clark HH, Murphy GL. 1982 Audience design in meaning and reference. *Adv. Psychol.* **9**, 287–299.
 113. Brennan SE, Clark HH. 1996 Conceptual pacts and lexical choice in conversation. *J. Exp. Psychol.: Learn. Mem. Cogn.* **22**, 1482–1493. (doi:10.1037/0278-7393.22.6.1482)
 114. Bromme R, Jucks R, Wagner T. 2005 How to refer to 'diabetes'? Language in online health advice. *Appl. Cogn. Psychol.* **19**, 569–586. (doi:10.1002/acp.1099)
 115. Heller D, Gorman KS, Tanenhaus MK. 2012 To name or to describe: shared knowledge affects referential form. *Topics Cogn. Sci.* **4**, 290–305. (doi:10.1111/j.1756-8765.2012.01182.x)
 116. Horton WS, Gerrig RJ. 2002 Speakers' experiences and audience design: knowing when and knowing how to adjust utterances to addressees. *J. Mem. Lang.* **47**, 589–606. (doi:10.1016/S0749-596X(02)00019-0)
 117. Horton WS, Gerrig RJ. 2005 The impact of memory demands on audience design during language

- production. *Cognition* **96**, 127–142. (doi:10.1016/j.cognition.2004.07.001)
118. Isaacs EA, Clark HH. 1987 References in conversation between experts and novices. *J. Exp. Psychol.: Gen.* **116**, 26–37. (doi:10.1037/0096-3445.116.1.26)
119. Schober MF, Clark HH. 1989 Understanding by addressees and overhearers. *Cognit. Psychol.* **21**, 211–232. (doi:10.1016/0010-0285(89)90008-X)
120. Yoon SO, Brown-Schmidt S. 2018 Aim low: mechanisms of audience design in multiparty conversation. *Discourse Process.* **55**, 566–592. (doi:10.1080/0163853X.2017.1286225)
121. Lockridge CB, Brennan SE. 2002 Addressees' needs influence speakers' early syntactic choices. *Psychon. Bull. Rev.* **9**, 550–557. (doi:10.3758/BF03196312)
122. Kuhlen AK. 2011 Assessing and accommodating addressees' needs: the role of speakers' prior expectations and addressees' needs. *Diss. Abstr. Int.: B: Sci. Eng.* **71**, 7128.
123. Galati A, Brennan SE. 2010 Attenuating information in spoken communication: for the speaker, or for the addressee? *J. Mem. Lang.* **62**, 35–51. (doi:10.1016/j.jml.2009.09.002)
124. Hagoort P, Levinson SC. 2014 Neuropragmatics. In *The cognitive neurosciences* (eds MS Gazzaniga, GR Mangun), pp. 667–674. Cambridge, MA: MIT Press.
125. Willems RM, de Boer M, de Ruiter JP, Noordzij ML, Hagoort P, Toni I. 2010 A dissociation between linguistic and communicative abilities in the human brain. *Psychol. Sci.* **21**, 8–14. (doi:10.1177/0956797609355563)
126. Kuhlen AK, Bogler C, Brennan SE, Haynes J-D. 2017 Brains in dialogue: decoding neural preparation of speaking to a conversational partner. *Soc. Cogn. Affect. Neurosci.* **12**, 871–880. (doi:10.1093/scan/nx018)
127. Haynes J-D, Rees G. 2006 Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* **7**, 523–534. (doi:10.1038/nrn1931)
128. Haynes J-D, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE. 2007 Reading hidden intentions in the human brain. *Curr. Biol.* **17**, 323–328. (doi:10.1016/j.cub.2006.11.072)
129. Bode S, Haynes J-D. 2009 Decoding sequential stages of task preparation in the human brain. *Neuroimage* **45**, 606–613. (doi:10.1016/j.neuroimage.2008.11.031)
130. Stolk A, D'Imperio D, di Pellegrino G, Toni I. 2015 Altered communicative decisions following ventromedial prefrontal lesions. *Curr. Biol.* **25**, 1469–1474. (doi:10.1016/j.cub.2015.03.057)
131. Boux I, Tomasello R, Grisoni L, Pulvermüller F. 2021 Brain signatures predict communicative function of speech production in interaction. *Cortex* **135**, 127–145. (doi:10.1016/j.cortex.2020.11.008)
132. Vanlangendonck F, Willems RM, Menenti L, Hagoort P. 2016 An early influence of common ground during speech planning. *Lang. Cogn. Neurosci.* **31**, 741–750. (doi:10.1080/23273798.2016.1148747)
133. Barr DJ, Keysar B. 2002 Anchoring comprehension in linguistic precedents. *J. Mem. Lang.* **46**, 391–418. (doi:10.1006/jmla.2001.2815)
134. Hanna JE, Tanenhaus MK, Trueswell JC. 2003 The effects of common ground and perspective on domains of referential interpretation. *J. Mem. Lang.* **49**, 43–61. (doi:10.1016/S0749-596X(03)00022-6)
135. Metzger C, Brennan SE. 2003 When conceptual pacts are broken: partner-specific effects on the comprehension of referring expressions. *J. Mem. Lang.* **49**, 201–213. (doi:10.1016/S0749-596X(03)00028-7)
136. Bard EG, Anderson AH, Sotillo C, Aylett M, Doherty-Sneddon G, Newlands A. 2000 Controlling the intelligibility of referring expressions in dialogue. *J. Mem. Lang.* **42**, 1–22. (doi:10.1006/jmla.1999.2667)
137. Barr DJ, Keysar B. 2005 Making sense of how we make sense: the paradox of egocentrism in language use. In *Figurative language comprehension: social and cultural influences* (eds HL Colston, AN Katz), pp. 21–41. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
138. Keysar B, Barr DJ, Balin JA, Paek TS. 1998 Definite reference and mutual knowledge: process models of common ground in comprehension. *J. Mem. Lang.* **39**, 1–20. (doi:10.1006/jmla.1998.2563)
139. Keysar B, Barr DJ, Balin JA, Brauner JS. 2000 Taking perspective in conversation: the role of mutual knowledge in comprehension. *Psychol. Sci.* **11**, 32–38. (doi:10.1111/1467-9280.00211)
140. Kronmüller E, Barr DJ. 2007 Perspective-free pragmatics: broken precedents and the recovery-from-preemption hypothesis. *J. Mem. Lang.* **56**, 436–455. (doi:10.1016/j.jml.2006.05.002)
141. Bögels S, Barr DJ, Garrod S, Kessler K. 2015 Conversational interaction in the scanner: mentalizing during language processing as revealed by MEG. *Cereb. Cortex* **25**, 3219–34. (doi:10.1093/cercor/bhu116)
142. Apperly IA, Samson D, Carroll N, Hussain S, Humphreys G. 2006 Intact first- and second-order false belief reasoning in a patient with severely impaired grammar. *Soc. Neurosci.* **1**, 334–348. (doi:10.1080/17470910601038693)
143. Dronkers NF, Ludy CA, Redfern BB. 1998 Pragmatics in the absence of verbal language: descriptions of a severe aphasic and a language-deprived adult. *J. Neurolinguistics* **11**, 179–190. (doi:10.1016/S0911-6044(98)00012-8)
144. Varley R, Siegal M. 2000 Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. *Curr. Biol.* **10**, 723–726. (doi:10.1016/S0960-9822(00)00538-8)
145. Varley R, Siegal M, Want SC. 2001 Severe impairment in grammar does not preclude theory of mind. *Neurocase* **7**, 489–493.
146. Willems RM, Benn Y, Hagoort P, Toni I, Varley R. 2011 Communicating without a functioning language system: implications for the role of language in mentalizing. *Neuropsychologia* **49**, 3130–3135. (doi:10.1016/j.neuropsychologia.2011.07.023)
147. Åsberg J. 2010 Patterns of language and discourse comprehension skills in school-aged children with autism spectrum disorders: comprehension in ASD. *Scand. J. Psychol.* **51**, 534–539. (doi:10.1111/j.1467-9450.2010.00822.x)
148. Diehl JJ, Benetto L, Young EC. 2006 Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *J. Abnorm. Child Psychol.* **34**, 83–98. (doi:10.1007/s10802-005-9003-x)
149. Lord C, Paul R. 1997 Language and communication in autism. In *Handbook of autism and pervasive developmental disorders* (eds D Cohen, F Volkmar), pp. 195–225, 2nd ed. New York, NY: Wiley.
150. Tager-Flusberg H. 2006 Defining language phenotypes in autism. *Clin. Neurosci. Res.* **6**, 219–224. (doi:10.1016/j.cnr.2006.06.007)
151. Terzi A, Marinis T, Francis K. 2016 The interface of syntax with pragmatics and prosody in children with autism spectrum disorders. *J. Autism Dev. Disord.* **46**, 2692–2706. (doi:10.1007/s10803-016-2811-8)
152. Wilkinson KM. 1998 Profiles of language and communication skills in autism. *Ment. Retard. Dev. Disabil. Res. Rev.* **4**, 73–79. (doi:10.1002/(SICI)1098-2779(1998)4:2<73::AID-MRDD3>3.0.CO;2-Y)
153. Deen B, Koldewyn K, Kanwisher N, Saxe R. 2015 Functional organization of social perception and cognition in the superior temporal sulcus. *Cereb. Cortex* **25**, 4596–4609. (doi:10.1093/cercor/bhv111)
154. Paunov AM, Blank IA, Fedorenko E. 2019 Functionally distinct language and Theory of Mind networks are synchronized at rest and during language comprehension. *J. Neurophysiol.* **121**, 1244–1265. (doi:10.1152/jn.00619.2018)
155. Cai ZG, Gilbert RA, Davis MH, Gaskell MG, Farrar L, Adler S, Rodd JM. 2017 Accent modulates access to word meaning: evidence for a speaker-model account of spoken word recognition. *Cognit. Psychol.* **98**, 73–101. (doi:10.1016/j.cogpsych.2017.08.003)
156. Van Berkum JJA, Holleman B, Nieuwland M, Otten M, Murre J. 2009 Right or wrong?: the brain's fast response to morally objectionable statements. *Psychol. Sci.* **20**, 1092–1099. (doi:10.1111/j.1467-9280.2009.02411.x)
157. Xu J, Abdel Rahman R, Sommer W. 2021 Who speaks next? Adaptations to speaker identity in processing spoken sentences. *Psychophysiology* **59**, e13948. (doi:10.1111/psyp.13948)
158. Ryskin R, Ng S, Mimnaugh K, Brown-Schmidt S, Federmeier KD. 2020 Talker-specific predictions during language processing. *Lang. Cogn. Neurosci.* **35**, 797–812. (doi:10.1080/23273798.2019.1630654)
159. Abdel Rahman R, Melinger A. 2011 The dynamic microstructure of speech production: semantic interference built on the fly. *J. Exp. Psychol.: Learn. Mem. Cogn.* **37**, 149–161. (doi:10.1037/a0021208)
160. Lin H-P, Kuhlen AK, Abdel Rahman R. 2021 Ad-hoc thematic relations form through communication: effects on lexical-semantic processing during language production. *Lang. Cogn. Neurosci.* **36**, 1057–1075. (doi:10.1080/23273798.2021.1900580)
161. Mädebach A, Kurtz F, Schriefers H, Jescheniak JD. 2019 Pragmatic constraints do not prevent the co-

- activation of alternative names: evidence from sequential naming tasks with one and two speakers. *Lang. Cogn. Neurosci.* **35**, 1073–1088. (doi:10.1080/23273798.2020.1727539)
162. Dell GS. 1986 A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* **93**, 283–321. (doi:10.1037/0033-295X.93.3.283)
163. Dell GS, Martin N, Schwartz MF. 2007 A case-series test of the interactive two-step model of lexical access: predicting word repetition from picture naming. *J. Mem. Lang.* **56**, 490–520. (doi:10.1016/j.jml.2006.05.007)
164. Dell GS, O'Seaghdha PG. 1991 Mediated and convergent lexical priming in language production: a comment on Levelt *et al.* (1991). *Psychol. Rev.* **98**, 604–614. (doi:10.1037/0033-295X.98.4.604)
165. Dell GS, O'Seaghdha PG. 1992 Stages of lexical access in language production. *Cognition* **42**, 287–314. (doi:10.1016/0010-0277(92)90046-K)
166. Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. 1997 Lexical access in aphasic and nonaphasic speakers. *Psychol. Rev.* **104**, 801–838. (doi:10.1037/0033-295X.104.4.801)
167. Levelt WJM, Schriefers H, Vorberg D, Meyer AS, Pechmann T, Havinga J. 1991 The time course of lexical access in speech production: a study of picture naming. *Psychol. Rev.* **98**, 122–142. (doi:10.1037/0033-295X.98.1.122)
168. Levelt WJM, Roelofs A, Meyer AS. 1999 A theory of lexical access in speech production. *Behav. Brain Sci.* **22**, 1–38. (doi:10.1017/S0140525X99001776)
169. Rapp B, Goldrick M. 2000 Discreteness and interactivity in spoken word production. *Psychol. Rev.* **107**, 460–499. (doi:10.1037/0033-295X.107.3.460)
170. Roelofs A. 1992 A spreading-activation theory of lemma retrieval in speaking. *Cognition* **42**, 107–142. (doi:10.1016/0010-0277(92)90041-F)
171. Abdel Rahman R, Melinger A. 2009 Semantic context effects in language production: a swinging lexical network proposal and a review. *Lang. Cogn. Process.* **24**, 713–734. (doi:10.1080/01690960802597250)
172. Abdel Rahman R, Melinger A. 2019 Semantic processing during language production: an update of the swinging lexical network. *Lang. Cogn. Process.* **34**, 1176–1192. (doi:10.1080/23273798.2019.1599970)
173. Ferstl EC, Neumann J, Bogler C, von Cramon DY. 2008 The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum. Brain Mapp.* **29**, 581–593. (doi:10.1002/hbm.20422)
174. Hagoort P. 2019 The neurobiology of language beyond single-word processing. *Science* **366**, 55–58. (doi:10.1126/science.aax0289)
175. Hintz F, Meyer AS, Huettig F. 2016 Encouraging prediction during production facilitates subsequent comprehension: evidence from interleaved object naming in sentence context and sentence reading. *Q. J. Exp. Psychol.* **69**, 1056–1063. (doi:10.1080/17470218.2015.1131309)
176. Rueschemeyer SA, Gardner T, Stoner C. 2015 The Social N400 effect: how the presence of other listeners affects language comprehension. *Psychon. Bull. Rev.* **22**, 128–134. (doi:10.3758/s13423-014-0654-x)
177. Kim D, Hommel B. 2019 Social Cognition 2.0: toward mechanistic theorizing. *Front. Psychol.* **10**, 2643. (doi:10.3389/fpsyg.2019.02643)
178. Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C. 2012 Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn. Sci.* **16**, 114–121. (doi:10.1016/j.tics.2011.12.007)
179. Hari R, Henriksson L, Malinen S, Parkkonen L. 2015 Centrality of social interaction in human brain function. *Neuron* **88**, 181–193. (doi:10.1016/j.neuron.2015.09.022)
180. Redcay E, Schilbach L. 2019 Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nat. Rev. Neurosci.* **20**, 495–505. (doi:10.1038/s41583-019-0179-4)
181. Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K. 2013 Toward a second-person neuroscience. *Behav. Brain Sci.* **36**, 393–414. (doi:10.1017/S0140525X12000660)
182. Verga L, Kotz SA. 2019 Putting language back into ecological communication contexts. *Lang. Cogn. Neurosci.* **34**, 536–544. (doi:10.1080/23273798.2018.1506886)
183. Osborne-Crowley K. 2020 Social cognition in the real world: reconnecting the study of social cognition with social reality. *Rev. Gen. Psychol.* **24**, 144–158. (doi:10.1177/1089268020906483)