Alessandra Chiera

Roma Tre University
ale.chiera@gmail.com

Alignment as the Sensorimotor Basis of the Evolution of Conversation

Abstract. Recent considerations on the nature of language recognize conversation as the central unit of analysis. Some approaches give a definition of conversation as parallel with that of cooperative action, with conversational success taking place when individuals converge to achieve a common goal. The present challenge of psycholinguistic is identifying the cognitive mechanisms underlying the process of convergence in conversation. Among these, interactional alignment, i.e. accommodation at many levels in dialogue, has received growing attention.

In this paper, I will propose that, when it comes to the evolutionary issue, alignment might be considered a good candidate to account for some early strategies used by individuals keen to communicate in the absence of a full-fledged code. The focus on mechanisms of low-level alignment paves the way for a sensorimotor and protoconversational account of language evolution.

Keywords: language evolution; conversation; alignment; sensorimotor cognition; pragmatics.

From Monologue to Conversation

Conversation is the cradle of language. Making such a claim does not entail facing any particular objections or controversies: that face-to-face conversation represents the most natural, basic and universal setting for language use is widely recognized (see Clark, 1996). It is how conversation works and what kind of psychological mechanisms implies that are at issue

in the literature. Traditionally, two diametrically opposed approaches are in place. On the one hand, many scholars from different lines of investigation inspired by the social sciences have considered language use as an entirely social process autonomous from the individual minds (e.g. Fishman, 1971). On the other hand, in several disciplines within the cognitive sciences, conversation has been tackled as an individual process involving a speaker and a listener that act as autonomous agents dealing with the production and reception of sentences in speech (e.g. Pickering et al., 2001).

If the former paradigm has completely reduced language to a cultural practice independent from the biology of individuals, it should be pointed out that the latter approach – within which I move – has traditionally considered language as an "internalist" matter. The "internalist" conception of language goes hand in hand with the idea that its minimum unit of analysis is the sentence. The centrality of the sentence relates with the prevailing model of the mind assumed by the cognitive scientists, i.e. the computational and modularist account by Fodor (1975, 1983), that, in turn, binds to the reference model of language, i.e. the Chomskyan Universal Grammar (e.g. Chomsky, 1995). As such models explain the construction of language meaning in terms of syntactic combination of mental symbols to process propositional structures – a procedure entirely guided by domain-specific systems having access only to strict linguistic content and independent from other cognitive domains –, the sentence is the core unit of language. This perspective affects the way of considering conversation: since meaning is the product of the atomic elements of the sentence, conversation is the result of a concatenation of sentences. In other words, conversation is the sum of two autonomous activities – encoding and decoding of strings of symbols – that do not involve a relation to context and social dynamics (except in marginal circumstances, negligible for a strict theory of language) (Cappelen & Lepore, 2005).

Despite its popularity resulting from the fact that it represents the most solid and empirically founded account of a certain conception of language (for a discussion, see Ferretti et al., 2018), the classical syntax-centric account appears to be an implausible model of conversation. To support this claim are theoretical and experimental considerations stressing that language use goes beyond the analysis of sentence constituent structure (e.g. Clark, 1996; Cosentino et al., 2013; Marini et al., 2014). Face-to-face conversation seems to be a deeply social activity implying the elaboration of several elements of the extra-linguistic context as conversation unfolds (e.g. Kim, 2014; van Berkum et al., 2008). Further, the success of conversation appears to be strictly tied to speaker's ability to adjust expressions on the recipient's point of view. In this regard, it has been emphasized that conversation requires

the elaboration of situation models of the whole conversational context (Zwaan & Radvansky, 1998), i.e., a mental representation of the main aspects of the situation. A conception of this kind implies a very different perspective compared to the context-free syntactic rules brought into play by the traditional account.

From these considerations follows that the "internalist" conception of language offered by the classical cognitive sciences has to be revised in order to provide a comprehensive account of language that includes conversation. Against the background of the second-generation cognitive sciences, conversation can be studied as the result of both individual and social aspects. Such an approach may provide profitable implications for a theory of language functioning and evolution. The first move in this direction is changing perspective on what is meant by language.

Conversation as Joint Action

Starting from the assumption that conversation implies more than the encoding and decoding of sentences, some scholars (e.g. de Ruiter et al., 2010; Noordzij et al., 2010) have focused on the interactional aspect underlying communication by isolating it from the linguistic system. The aim was to investigate the strategies employed to converge on meaning without sharing a common code. The Tacit Communication Game (TCG) used to this end required participants to engage in a dialogic coordination task trying to realize a goal configuration using shapes on a grid. The results interestingly showed that the difficulty of the communication problem weighed equitably on both the sender and the receiver and that success was affected by feedback from receiver to sender. This data suggests that the functional context structures conversational interaction, by driving joint decisions. Further, it was observed a cerebral overlap of brain activity in the sender and in the listener during the planning of the communicative action by the former and the recognition of that action by the latter. The result is consistent with an fMRI study by Stephens and colleagues (2010) including participants engaged in attending to the speech generated by a speaker. They found that the speaker's and the listeners' brain activity were coupled both at the spatial and temporal level. Interestingly, some areas in the listener's brain were activated before they were in the speaker's, suggesting forms of anticipatory processing necessary for coordination. Overall, these findings corroborate the idea that the nature of conversation has to be investigated looking at the actions performed by individuals who interact in various use-contexts to build portions of meaning

(Holtgraves, 2001). To this extent, conversation can be interpreted as a form of joint action, in which people act together to achieve a common goal (Clark, 1996, 2002).

A similar redefinition of the nature of language in relation to its interactive dimension lies within the unique paradigm alternative to the Chomskyan model: the ostensive-inferential model of communication (OIMC) or relevance theory (Sperber & Wilson, 1986/1995). The OIMC establishes a shift from the study of language in abstract terms to the investigation of language in interaction. According to the model, corroborated by solid cognitive data, the thesis of a decontextualized language is unfounded since language is always constrained by its use and involves pragmatic non-linguistic phenomena necessary to recover what is communicated from what is said (Carston, 2002). Specifically, this recovery is oriented by the speaker's meaning, i.e. the communicative intention aimed to achieve a certain effect on the hearer's mind, and the listener's recognition of that intention. In this view, the entire communication process is a collaborative practice grounded on the interlaced work of speaker's recipient design – the tailoring of messages for specific addressees – and listener's intention recognition. This conception centered on a definition of language in terms of use matches with the idea that the social-cognitive infrastructure underlying engagement with other minds – a mindreading system – represents a major crux in the investigation of language processing as well as its origins and evolution (e.g. Dunbar, 1998; Origgi & Sperber, 2000; Scott-Phillips, 2014; Tomasello, 2008). In fact, the focus on the role of mindreading as a precondition for the phylogenetic development of language is at the heart of much current work (e.g. Scott-Phillips, 2014; Origgi & Sperber, 2000). Before considering this issue within the evolutionary framework, I intend to focus on specific mechanisms involved in the pragmatic foundation of language. Such mechanisms are at the root of mindreading abilities and are specifically involved in the most central feature of all joint actions, namely coordination, the mutual process by which interlocutors build a common ground to converge on meaning (Gambi & Pickering, 2011).

Alignment in Conversation

Recent investigations define conversation as an interpersonal synergy (Fusaroli et al., 2014) constrained by low-level coordinative mechanisms – that is to say, forms of coordination are a pre-requisite for the collaborative

dimension of conversation to emerge and work (Levinson, 2006). Coordination in conversation may concern convergence on both content – the shared goal to achieve – and processes – the psychological systems recruited in the joint action (Guardiola & Bertrand, 2013).

The most systematic theory providing a unified account of several aspects of coordinative phenomena in conversation has been developed by Pickering and Garrod (2004). According to their interactive-alignment theory, conversation is guided by the alignment of linguistic behaviors between interactants. Notwithstanding a symmetry between linguistic sequences is not always observed in conversation, Pickering and Garrod highlight that when a convergence of language structures at multiple levels – i.e., phonetic, syntactic, semantic and so on – is automatically established, this increases and facilitates interaction (Fusaroli & Tylén, 2012). The notion of alignment represents a powerful process able to account for the effortless character of conversational interaction: the underlying mechanism is, in fact, that of structural priming that works within the speaker's production system and the hearer's comprehension system (Pickering & Ferreira, 2008). How does it function? Structural priming implies the mutual entrainment of individuals' representational states via a progressive synchronization of linguistic sequences: alignment at any level will spread to higher levels leading ultimately to the alignment of situation models, i.e., the conceptual models built to catch the main aspects of a contextual situation. In other words, individuals prime each other at many linguistic levels and, because of a sort of ripple effect, they achieve a common comprehension of the situation under discussion. The theoretical assumption of the interactive-alignment model is that speaker and listener share the same representations, namely they are guided by a representational parity between production and comprehension. This assumption has been widely confirmed by neuroimaging investigations (e.g. Menenti et al., 2011; Stephens et al., 2010) that highlight a neural basis of alignment of mental states during production and comprehension processes.

Explaining conversation in terms of imitation-like alignment of linguistic structures is a way to propose a mechanistic model of language interaction based on "simple" mechanisms that enable an unconscious and rapid comprehension process between interlocutors. To this extent, the interactive-alignment account provides a plausible hypothesis alternative to the classical cognitivist theory of language centered on an autonomous information transfer account of monologue. The focus on low-level mechanisms makes the mechanistic account of conversation a refined model that could be

useful also in an evolutionary perspective. The next paragraph deals with the implications of this model for the issue of the origins and evolution of language.

The Evolutionary Issue: What the Mechanistic Model Explains and What It Does Not

Over the last few years, some approaches to the evolution of language have acknowledged conversation, considered in a broad sense in terms of a medium of social interaction, as the phylogenetically primary form of language behavior (Wacewicz et al., 2017). Depending on the starting framework, the nature of conversation in the first stages of language evolution has been framed in connection to behaviors such as grooming (Dunbar, 1996; MacNeilage, 2008) or intra-group cooperative dynamics (Gärdenfors, 2003).

On account of its attention towards basic mechanisms able to explain complex conversational processes by observing a principle of cognitive parsimony, the mechanistic model of conversation has been recently used to illuminate the evolutionary debate. Specifically, Pickering and Garrod (2017) have put forward the idea that interactive alignment can be considered a powerful means of developing an evolutionary motivated account of how language changes and is transmitted. The hypothesis is that alignment might have guided automatic transmission during usage through a routinization process. As the authors highlight, in fact, conversational interaction is very repetitive and seems to enable routines, namely stored representations in Jackendoff's (2002) terms: linguistic information that is stored as a lexical representation mapping between phonological, syntactic and conceptual aspects. Pickering and Garrod suggest that, if long-term routines become aligned in a population, group alignment may occur leading to systematic routinization across the community. In this regard, automatic community alignment can be viewed as a mechanism for driving language change.

Although the proposal appears fascinating, it is my contention that it has two important shortcomings: the first regards some concerns raised about the cognitive plausibility of the mechanistic account as a model of conversational processing that, in turn, have important consequences on the evolutionary plane; the second limit relates to the question of what the interactive-alignment model can explain and what it cannot from the specific evolutionary point of view.

As for the first limitation, some criticisms have been moved to the mechanistic theory as a comprehensive model of language functioning.

Within a certain tradition, the increase of investigations in structural priming has made it the prevailing mechanism of accounting for the processes involved in conversation. However, the assumption of a progressive automatic alignment toward systematic synchronization of each other's linguistic sequences is controversial: observations on interactive alignment show that it may not be automatic and that it is context-sensitive depending on factors like group identity and affiliation (see van Baaren et al., 2009). From this point of view, considering alignment in rigid terms as proposed by Pickering and Garrod (2004) does not entirely account for the success of conversation (Fusaroli et al., 2014). In addition, and more interestingly, the interactive-alignment model appears unable to explain the actual collaborative nature of conversation as it states. In fact, the exclusive reference to an automatic priming mechanism for explaining the entire process of convergence toward a final coordination of situation models seems not to catch the collaborative enterprise realized by interactants to modulate and keep track of each other's intentions. If communication is a joint practice oriented by the speaker's recipient design and listener's intention recognition where interlocutors should coordinate on high-level behaviors such as establishing common ground and achieving the goal of the conversation, structural priming implying the mutual accommodation of mental states via an automatic alignment of linguistic structures appears to be an insufficient mechanism. What seems to be missing is a level of analysis explaining how one can move from linguistic alignment to convergence at the level of conceptual models. In fact, it is very controversial claiming that from accommodation in linguistic representations follows accommodation of situation models. As stated by the OIMC, linguistic utterances provide just indirect hints of speaker's situation model; the listener uses linguistic information to reconstruct that model starting from the evidence (Branigan, 2004). Consistent with this idea, it has been showed that alignment at any linguistic level may occur along with disaligned or misaligned situation models between individuals both in typical and clinical populations (Ferreira et al., 2008; Garrod & Clark, 1993; Stewart et al., 2008). To this extent, automatic mechanisms involved in the possibility that some linguistic structure are selected by speaker and hearer in conversation is not a sufficient condition for the intentional and conceptual alignment that Pickering and Garrod (2004) intend to explain (Schiller & de Ruiter, 2004).

This concern looks particularly important from an evolutionary perspective and opens the way to the second question suggested above about what the interactive-alignment model can and what it cannot explain as for the topic of language evolution. As discussed above, Pickering and Garrod

(2017) refer to a mechanism of group alignment to explain how language can change in conversation through a routinization process. Despite the fact that the reference to a mechanism of this kind looks suitable in an evolutionary perspective, it may account for some restricted aspects of language evolution, namely for the automatic and rapid transmission of certain linguistic structures during language use when conversation is already linguistically organized. From this point of view, the interactive-alignment model takes place against a background of several theories of language evolution that focus attention almost completely on the analysis of strictly linguistic components (e.g. Christiansen & Chater, 2008; Christiansen & Kirby, 2003; Fitch, 2010). As a result, the model turns out to present the same limitation of those theories: that of falling into the fallacy of taking for granted from the beginning something that must be explained (Tomasello, 2008). In fact, from an evolutionary point of view, considering the rise of human communication starting from a symbolic code means assuming a preexisting form of communication which is merely encoded (Chiera, 2016). This result is even paradoxical if we consider that the interactive-alignment model arises precisely to construct a conception of language based on its foundation that is identified in the more general ability to process joint actions. However, focusing on the convergence of linguistic structures as a sufficient mechanism for the comprehension process, the interactive-alignment model has no explanation for the pragmatic link between the complementary actions constantly negotiated by speaker and hearer to converge on the intentional level. An explanation of such a link is crucial for a theory of language origins and evolution since at the beginning, when codes were not yet full-formed, just pragmatic abilities of convergence on a representation of common ground – a representation of other's actions and intentions alongside one's own – might have helped interactants to keep alive the early forms of communicative interaction (Origgi & Sperber, 2000).

Starting from these considerations, it is my claim that language evolution has to be analyzed in regard to broader capacities that are in charge of the ability to establish a common ground with the interlocutor. This pragmatic perspective on language evolution can be profitably combined with a protoconversational hypothesis, namely with the idea that the pragmatic processes involved in the construction of a shared conceptual space characterize the early stages of language evolution in terms of a dynamic exchange, situated within a jointly determined – and constantly evolving – action context interactively negotiated (Chiera, 2016). What abilities were involved in this scenario? Alignment may be considered an effective means in evolutionary key as a low-level mechanism that could have fostered

linguistic communication. However, because of the criticisms outlined above, in my perspective the focus on alignment is related to aspects other than those discussed by Pickering and Garrod (2004, 2017). A cognitive architecture for coordination of interactions in protoconversation should comprise a set of mechanisms underlying mutual understanding between interlocutors at both a lower level, i.e. bodily actions, and a higher level, i.e. intentions and situation models (Knoblich et al., 2011). My hypothesis is that the low-level bodily aspects – the roots of mindreading abilities – might have structured protoconversation before linguistic conversation.

Low-Level Alignment in Protoconversation

In conversation, coordination concerns both high-level and low-level aspects: if the higher-level relates to the coordination of meaning at the representational plane, the lower-level consists in nonverbal synchronization of behaviors, i.e, postures, bodily movements, facial expressions (Latif et al., 2014). This lower level of coordination, in turn, can involve convergence of aspects related to form and timing: in Wacewicz et al. (2017) we referred to the former in terms of mimicry and to the latter in terms of synchrony. Whereas mimicry concerns the form of coordinated behavioral patterns mutually adopted by interactants, synchrony refers to the timing of individual actions in a collaborative context, e.g. turn-taking behaviors.

In the present paper, I focus on the mimicry mechanisms as particularly interesting in evolutionary key. In fact, a promising way of identifying the roots of language is to explore the hypothesis that low-level forms of coordination might have represented an evolutionary basis for the cooperative dimension of language by guiding the construction of higher levels, namely the representational and intentional ones. Such a hypothesis dovetails with an approach to conversation as an interactive activity requiring joint-construction in terms of bodily mimicry more than linguistic convergence (Guardiola & Bertrand, 2013). According to this approach, reciprocal behavioral and physiological mimicry (Chartrand & Bargh, 1999; Dijksterhuis & Bargh, 2001) consisting in the involuntary adjustment of the form of one's movements to that of the interlocutor is part of the comprehension process. In this regard, individuals engaged in conversation have been observed to unconsciously and spontaneously make their expressions, gestures, body posture and gaze patterns more similar over time (Dijksterhuis & Bargh, 2001; Richardson & Dale, 2005). Interestingly, low-level coordination correlates with mutual understanding and has been hypothesized to embody the goals of the joint

activity (Shockley et al., 2009). Thus, rather than representing a secondary aspect, these forms of basic alignment appear to be significant for the success of conversation. An important role is that of facilitating focused interaction by emphasizing the processes the speaker employs to mark her discourse and the sustained attention of the interactants (see Wacewicz et al., 2017). To this extent, low-level coordinative processes might provide a cooperative foundation in face-to-face conversation. In this regard, the literature suggests that coordination at the level of mimicry affects the degree of cooperation between interlocutors. Shockley et al. (2003) reported that bodily alignment in conversation promotes cognitive coordination thus leading the success of cooperative communication. Similarly, Richardson et al. (2007), measuring the bodily attunement of individuals solving a cooperative "puzzle task", found that eye movement coordination is sensitive to cognitive factors: the gaze patterns of the participants were more coupled when they had listened to the same background information than when they had not. These considerations suggest that reciprocal context-sensitive alignment may play a crucial role in the processing of ordinary conversation.

The focus on low-level coordination looks even more interesting in an evolutionary perspective since it might represent the foundation of the progressive emergence of cooperative communication. In the attempt to identify the prerequisites of cooperative communication within a comparative perspective, some have observed that many nonhuman animals display coordinative behaviors similar to those discussed above. If mimicry in a broad sense has been showed to belong to a wide range of animals, e.g. flocking birds, schooling fish and dolphins (de Waal, 2009), interestingly complex forms of mimicry, including affect coordination, are displayed by both monkeys and nonhuman apes. These species are able to follow the gaze direction of their conspecifics and to progressively adjust their attention when the latter re-orient their body posture (Kaminski et al., 2004), to be distressed by the distress of a conspecific and attempt to eliminate his pain (Preston & de Waal, 2002) and to engage in contagious laugh-like vocalizations (Provine, 2000).

My main idea is that the cooperative nature of human language rests on a scaffolding of similar non-representational mechanisms, which in the human evolutionary history might have progressively led to convergence on the higher level (Wacewicz et al., 2017). In fact, although automatic and unconscious these low-level mechanisms provide some hints about the other's mental states. This assumption is consistent with the pragmatic perspective adopted by the OIMC. Sperber and Origgi (2010) have suggested that it is plausible to conceive the early stages of communication as characterized

by a fragmentary and ambiguous code to be interpreted by means of the ability to attribute mental states to others. I agree that at the origin the code played a little role and that the ability to make predictions of the speaker's meaning was the crucial point. However, unlike Sperber and Origgi account in which a great attention is given to sophisticated abilities of mindreading requiring high-level cognitive skills, my proposal is that the construction of a shared mental space involved in building joint communicative actions may be ascribable to low-level sensorimotor mechanisms. This idea is more consistent with the evolutionary conception and with the data documenting that unconscious, rapid and automatic behavioral coordination may facilitate predictive strategies and then comprehension (Knoblich & Jordan, 2003). Within the second-generation cognitive sciences, a neurocognitive paradigm assuming the principles of evolutionary theory as points of reference provides convincing hypotheses on the nature of the abilities we identified as crucial for the emergence of conversation: the embodied theories of mind, which claim that supposedly high processes are grounded in low sensory, motor and emotional processes (e.g. Barsalou, 1999; Kemmerer, 2010; Gallese & Lakoff, 2005).

A Sensorimotor Hypothesis of Language Evolution

The central tenet of the embodied perspective is that cognition is strictly anchored to experience taking place in a real-world environment (Wilson, 2002). Since cognitive processes are mainly designed to guide actions in this environment, the mechanisms of perception and action have a primacy in the explanation of even sophisticated abilities. The shift of paradigm promoted by embodied cognition has deeply influenced theories of language functioning, leading to the idea that language processing is founded on the multimodal simulation of perceptions, actions and emotions (e.g. Barsalou, 1999; Pecher & Zwaan, 2005). Such an idea has important implications for the issue of language origins and evolution as well. The focus on sensorimotor mechanisms as the roots of linguistic phenomena paves the way to a truly Darwinian perspective centered on the idea that language can be reframed in terms of a constituent contribution provided by low processes intrinsically linked to the domain of action (see Ferretti et al., 2018). Although the notion of perception-action has been used in different ways to account for several aspects, for what we are concerned here this notion is relevant to language origin and evolution for a main reason: the perception-action link might have been exploited by language in the context of protoconversational interactions

providing a mechanism for reading other's intentions on the basis of the understanding of their actions (Blakemore & Decety, 2001). This claim, at the heart of the motor theory of social cognition, emphasizes a pragmatic approach to language evolution which is in line with the proposal of this paper.

The main hypothesis in the present article is, in fact, that low-level coordinative phenomena taking place unconsciously and automatically, i.e. mimicry, might have facilitated protoconversational interactions promoting coordination in the early stages of communication. Specifically, low-level mechanisms of mimicry might have played a crucial role in orienting processes of pragmatic alignment that allowed the coordination of what speaker meant and what the addressee understood him to mean in the context of a fragmentary sharing of code. The embodied perspective can account for similar dynamics. According to the motor account of social cognition, in fact, perception-action systems can serve a self-other matching function (e.g. Gallese, 2009) leading to the shaping of social interactions. Neurocognitive literature has provided strong evidence in support of these intersubjective bases of embodied cognition (e.g. Galantucci & Sebanz, 2009), showing that the capacity to understand other people's intentions is grounded on the observation of their actions (Ramenzoni et al., 2008; Richardson & Dale, 2005). The underlying assumption is that representing a perceived action activates a representation of the motor intention which would convey the social and communicative intention as well (Gallese, 2006; Iacoboni et al., 2005). In this perspective, building joint activities involves representing other's actions alongside one's own (Buccino et al., 2004) on the basis of a common vocabulary of actions and representations (Rizzolatti & Luppino, 2001). From this idea follows an important notion for the issue of language evolution, namely the notion of prediction. If the speaker can represent other people's actions alongside her own through an internal simulation of the observed action in her own motor system, this simulation can be used to formulate predictions about the other's behaviors by computing "forward models" (Wilson & Knoblich, 2005). The neuroscientific literature on joint actions shows that the use of other-generated actions to drive such predictions allows immediate comprehension (Blakemore & Decety, 2001; Buccino et al., 2004).

Overall, these considerations about the role of the perception-action link as a plausible basis of the coordinative processes involved in the construction of joint actions provide a key to interpret the early stages of communicative interactions. The mindreading abilities brought into play by Sperber and Origgi (2010) as necessary for a structured linguistic communication to

develop can be reframed against the background of the embodied conception in terms of more basic sensorimotor capacities. The motor account of social cognition offers indications to state that mirroring other people's actions is a way to directly comprehend the intentions of the actor who performed them (Goldman, 2006). To this extent, the sensorimotor system would represent the key to pragmatics and, in an evolutionary perspective, embodiment might have had a crucial role in meaning-making processes within the context of social action (Gilissen, 2005; Tettamanti et al., 2005). A similar hypothesis allows us to give a prominent role to mechanisms of bodily alignment realized by sensorimotor processes that might have represented a basis fostering successive forms of cooperation necessary for the construction of the conversational plane.

Conclusions

The purpose of the present paper was offering a hypothesis about the sensorimotor foundation of language functioning and evolution. Starting from the assumption that language use can be profitably defined in terms of a case of joint action, I focused on the specific role of a set of bodily mechanisms involved in the coordination of interactions, i.e. alignment. My claim was that a perception-action link underlying forms of pragmatic alignment might have served as a basis for prediction of others' actions that, in turn, might have fostered higher levels of alignment, i.e. intentional and conceptual levels, necessary for comprehension and for a shared code to develop. This hypothesis is consistent with the idea that low-level sensorimotor processes might have characterized the early forms of language use in protoconversational terms.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences*, 22(4), 577–660.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, 2(8), 561–567.

Branigan, H. P. (2004). Full alignment of some but not all representations in dialogue. *Behavioral & Brain Sciences*, 27(2), 191–192.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., ... Rizzolatti, G. (2004). Neural circuits involved in the recognition

of actions performed by nonconspecifics: An fMRI study. *Journal of Cognitive Neuroscience*, 16(1), 114–126.

Cappelen, H., & Lepore, E. (2005). *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell.

Carston, R. (2002). Linguistic meaning, communicated meaning and cognitive pragmatics. *Mind & Language*, 17(1–2), 127–148.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H. (2002). Speaking in time. *Speech Communication*, 36(1), 5–13.

Chartrand, T., & Bargh, J. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893–910.

Chiera, A. (2016). From conversation to language: An evolutionary sensory-motor account. In J. Zlatev, P. Konderak, & G. Sonnesson (Eds.), *Meaning, mind and communication: Explorations in cognitive semiotics* (pp. 189–198). Frankfurt am Main: Peter Lang.

Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(05), 489–509.

Christiansen, M. H., & Kirby, S. (Eds.). (2003). *Language evolution*. Oxford: Oxford University Press.

Cosentino, E., Adornetti, I., & Ferretti, F. (2013). Processing narrative coherence: Towards a top-down model of discourse. *Open Access Series in Informatics (OASICS)*. Special Issue *2013 Workshop on Computational Models of Narrative*, 32, 61–75.

de Ruiter, J. P., Noordzij, M. L., Newman-Norlund, S., Hagoort, P., Levinson, S. C., & Toni, I. (2010). Exploring the cognitive infrastructure of communication. *Interaction Studies*, 11(1), 51–77.

de Waal, F. (2009). *Age of empathy: Nature's lessons for a kinder society*. New York: Crown Publishing.

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–40). San Diego: Academic Press.

Dunbar, R. I. (1996). *Gossip, grooming, and the evolution of language*. Cambridge: Harvard University Press.

Dunbar, R. I. (1998). The social brain hypothesis. *Brain*, 9(10), 178–190.

Ferreira, V. S., Bock, K., Wilson, M., & Cohen, N. J. (2008). Memory for syntax despite amnesia. *Psychological Science*, 19(9), 940–946.

Ferretti, F., Adornetti, I., Chiera, A., Cosentino, E., & Nicchiarelli, S. (2018). Introduction: Origin and evolution of language – An interdisciplinary perspective. *TOPOI*, 37(2), 219–234.

Fishman, J. A. (Ed.). (1971). *Advances in the sociology of language use, Vol. 1: Basic concepts, theories, and problems: Alternative approaches*. The Hague: Mouton.

Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.

Fodor, J. A. (1975). *The language of thought*. New York: Thomas Crowell.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge: MIT Press.

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

Fusaroli, R., & Tylén, K. (2012). Carving language for social coordination: A dynamical approach. *Interaction Studies*, *13*(1), 103–124.

Galantucci, B., & Sebanz, N. (2009). Joint action: Current perspectives. *Topics in Cognitive Science*, 1(2), 255–259.

Gallese, V. (2006). Intentional attunement: A neurophysiological perspective on social cognition and its disruption in autism. *Brain Research*, 1079(1), 15–24.

Gallese, V. (2009). Mirror neurons, embodied simulation, and the neural basis of social identification. *Psychoanalitic Dialogues*, 19(5), 519–536.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3–4), 455–479.

Gambi, C., & Pickering, M. J. (2011). A cognitive architecture for the coordination of utterances. *Frontiers in Psychology*, 2, 1–14.

Garrod, S., & Clark, A. (1993). The development of dialogue co-ordination skills in schoolchildren. *Language and Cognitive Processes*, 8(1), 101–126.

Gilissen, E. (2005). Imitation systems, monkey vocalization, and the human language. *Behavioral & Brain Sciences*, 28(2), 133–134.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. New York: Oxford University Press.

Guardiola, M., & Bertrand, R. (2013). Interactional convergence in conversational storytelling: When reported speech is a cue of alignment and/or affiliation. *Frontiers in Psychology*, 4, 705.

Holtgraves, T. M. (2001). *Language as social action: Social psychology and language use*. Mahwah: Lawrence Erlbaum Associates.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, 3(3), e79.

Jackendoff, R. (2002). *Foundations of language*. Oxford: Oxford University Press.

Latif, N., Barbosa, A. V., Vatiokiotis-Bateson, E., Castelhano, M. S., & Munhall, K. G. (2014). Movement coordination during conversation. *PLoS One*, 9(8), e105036.

Kaminski, J., Call, J., & Tomasello, M. (2004). Body orientation and face orientation: Two factors controlling apes' requesting behavior from humans. *Animal Cognition*, 7(4), 216–223.

Kemmerer, D. (2010). How words capture visual experience. In B. Malt & P. Wolff (Eds.), *Words and the mind* (pp. 289–329). Oxford: Oxford University Press.

Kim, M. H. (2014). Why self-deprecating? Achieving 'oneness' in conversation. *Journal of Pragmatics*, 69, 82–98.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 59–101). Burlington: Academic Press.

Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), 1006–1016.

Levinson, S. C. (2006). On the human "interaction engine". In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Oxford: Berg.

MacNeilage, P. F. (2008). *The origin of speech*. Oxford: Oxford University Press.

Marini, A., Zettin, M., & Galetto, V. (2014). Cognitive correlates of narrative impairment in moderate traumatic brain injury. *Neuropsychologia*, 64, 282–288.

Menenti, L., Gierhan, S. M., Segaert, K., & Hagoort, P. (2011). Shared language: Overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*, 22(9), 1173–1182.

Noordzij, M. L., Newman-Norlund, S. E., de Ruiter, J. P., Hagoort, P., Levinson, S. C., & Toni, I. (2010). Neural correlates of intentional communication. *Frontiers in Neuroscience*, 4(188), 1–7.

Origgi, G., & Sperber, D. (2000). Evolution, communication and the proper function of language. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the human mind: Language, modularity and social cognition* (pp. 140–169). Cambridge: Cambridge University Press.

Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. New York: Cambridge University Press.

Pickering, M., Clifton, C., & Crocker, M. (2001). Architectures and mechanism in sentence comprehension. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and mechanism for language processing* (pp. 1–28). Cambridge: Cambridge University Press.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3), 427.

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

Pickering, M. J., & Garrod, S. (2017). Priming and language change. In M. Hundt, S. Mollin, & S. E. Pfenninger (Eds.), *The changing English language: Psycholinguistic perspectives* (pp. 173–190). Cambridge: Cambridge University Press.

Preston, S. D., & de Waal, F. B. M. (2002). The communication of emotions and the possibility of empathy in animals. In S. G. Post, L. G. Underwood,

J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruistic love: Science, philosophy, and religion in dialogue* (pp. 284–308). Oxford: Oxford University Press.

Provine, R. R. (2000). *Laughter: A scientific investigation*. New York: Viking.

Ramenzoni, V. N., Riley, M. A., Davis, T., Shockley, K., & Armstrong, R. (2008). Tuning in to another person's action capabilities: Perceiving maximal jumping-reach height from walking kinematics. *Journal of Experimental Psychology: Human Perception and Performance*, 34(4), 919–928.

Richardson, D., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Sciences*, 18(5), 407–413.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, 31(6), 889–901.

Schiller, N. O., & de Ruiter, J. P. (2004). Some notes on priming, alignment, and self-monitoring. *Behavioral & Brain Sciences*, 27(02), 208–209.

Scott-Phillips, T. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. London: Palgrave MacMillan.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 326–332.

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

Sperber, D., & Origgi, G. (2010). A pragmatic perspective on the evolution of language. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language: Biolinguistic perspectives* (pp. 124–131). Cambridge: Cambridge University Press.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford: Blackwell. (Original work published 1986)

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, 107(32), 14425–14430.

Stewart, S. L., Corcoran, R., & Drake, R. J. (2008). Alignment and theory of mind in schizophrenia. *Cognitive Neuropsychiatry*, 13(5), 431–448.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., [...] Perani D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17, 273–281.

Tomasello, M. (2008). *Origins of human communication*. Cambridge: MIT Press.

van Baaren, R., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1528), 2381–2389.

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

Wacewicz, S., Zywiczynski, P., & Chiera, A. (2017). An evolutionary approach to low-level conversational cooperation. *Language Sciences*, 63, 91–104.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9(4), 625–636.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3), 460–473.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185.