



The dialogically extended mind: Language as skilful intersubjective engagement

Action editor: Ron Sun

Riccardo Fusaroli^{a,b,*}, Nivedita Gangopadhyay^c, Kristian Tylén^{a,b}

^a Center for Semiotics, Institute for Aesthetics and Communication, Aarhus University, Building 1485, 6th Floor, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark

^b Interacting Minds, Aarhus University, Building 1483, 3rd Floor, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark

^c Center for Mind, Brain and Cognitive Evolution, Institute for Philosophy II, Ruhr-University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany

Abstract

A growing conceptual and empirical literature is advancing the idea that language extends our cognitive skills. One of the most influential positions holds that language – *qua* material symbols – facilitates individual thought processes by virtue of its material properties (Clark, 2006a). Extending upon this model, we argue that language enhances our cognitive capabilities in a much more radical way: the *skilful engagement* of public material symbols facilitates evolutionarily unprecedented modes of collective perception, action and reasoning (interpersonal synergies) creating *dialogically extended minds*. We relate our approach to other ideas about *collective minds* (Gallagher, 2011; Theiner, Allen, & Goldstone, 2010; Tollefsen, 2006) and review a number of empirical studies to identify the mechanisms enabling the constitution of interpersonal cognitive systems.

© 2013 Elsevier B.V. All rights reserved.

Keywords: Language and cognition; Language usage; Extended mind; Dialogue; Distributed cognition; Social interactions

1. Introduction

The intimate relation between language and cognition has been long recognised across a broad range of scientific and philosophical disciplines. However, the exact nature of the relation is still widely debated, cf. the different perspectives in A. Clark, 2006b, H.H. Clark, 1996; Fodor, 2008; Fusaroli, 2011; Geeraerts & Cuyckens, 2007; Tylén, Weed, Wallentin, Roepstorff, & Frith, 2010. One of the more recent developments considers language from the perspective of *active vehicle externalism*. In this perspective, language is regarded as an external culturally evolved tool that interacts with our biological cognitive systems facili-

tating and actively supporting certain cognitive processes (Clark, 2006a, 2006b). Language is thus portrayed as ‘a mind-transforming cognitive scaffolding: a persisting, though never stationary, symbolic edifice’ (Clark, 2008), which – thanks to its materiality and freedom from the immediate context – gives a more stable structure to thought. Internalist positions, for instance Fodor (2008), posit an innate language of thought to explain the symbolic structure of certain human cognitive processes. A. Clark on the contrary argues that it is the actual use of external material symbols – which in some cases can be internalized – that enables individual cognizers to think symbolically by constraining and focussing their perceptual and attentional strategies more effectively.

However, Clark’s position tends to neglect a crucial aspect of language, which radically extends its description as an instance of active vehicle externalism: language as a social activity. In most cases, humans do not simply engage

* Corresponding author at: Center for Semiotics, Institute for Aesthetics and Communication, Aarhus University, Building 1485, 6th Floor, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark. Tel.: +45 28890881.

E-mail addresses: fusaroli@gmail.com (R. Fusaroli), Nivedita.Gangopadhyay@rub.de (N. Gangopadhyay), kristian@cfin.dk (K. Tylén).

the material symbols *per se*, but employ them in interaction with other individuals, for instance in contexts of regulation of social relations, or coordination of complex actions and problem solving activities (cf. H.H. Clark, 1996; Brennan et al., 2010; Vygotsky, 1978; Habermas & Cooke, 1998; Fusaroli & Tylén, 2012; Pickering & Garrod, 2004; Hasson et al., 2012). Extending upon A. Clark's proposal, we stress how language enables *skilful intersubjective engagement*, that is the coordination of individual cognitive systems giving rise to composite units that exceed the capabilities of their parts (cf. the notion of *interpersonal synergies* (Riley et al., 2011; Fusaroli, Raczaszek-Leonardi, & Tylén, 2013). Rather than a simple cognition-enhancing external resource for individual cognition, language thus constitutes a new and evolutionarily unprecedented mode of socially extended cognition (Donald, 2001)¹. Linguistic activity is a means by which individuals come to jointly apprehend and manipulate information to create informational and behavioural interpersonal synergies, which potentially outstretch the cognitive abilities of any of the individuals were they on their own. Thus, language as a skilful intersubjective activity *de facto* constitutes dialogically extended minds.

We introduce our proposal by discussing A. Clark's idea of language as a tool. From Clark's perspective of active externalism, we argue that the bodily basis of language provides an initial step towards liberating linguistic meaning from the confines of purely internal neural processing, making language into something we *do*. However, since language use and development importantly anchors cognition into the social world, we will take the claim a step further: language is something we *do together*. We thus propose that language is a 'doubly-extended' cognitive phenomenon: not only is it robustly grounded in the agent's bodily engagement with the world, as hinted by Clark, but it also further extends this engagement into the social world through embodied social dynamics. We support this claim with reference to empirical findings on linguistic coordination, and point to possible mechanisms for the creation of interpersonal synergies. Finally we will discuss in which way our proposal complement other work on collective minds and respond to some possible critiques.

2. Language as tool for individual minds

2.1. Cognition beyond the boundaries of skull and skin

A. Clark's work introduces the notion of language as tool in order to challenge one of the fundamental assumptions in the contemporary philosophy of language and cognitive science, namely that innate internal linguistic representations are the necessary presupposition for the development and use of language as well as for human

¹ By "extended" we simply mean that the cognitive activity extends beyond the individual organism. However, we do not intend to engage here the debate of the primacy or not of the social world.

thought (Fodor, 1975, 2008). In opposition to such strongly internalistic conceptions, A. Clark develops the extended mind hypothesis: an active externalist conception in which not only internal neural structures, but also elements of the external world can constitute representational vehicles of mental states, that is, actively support the formation and storage of cognitive content. The idea is initially presented in the fictive case of the Alzheimer's patient Otto who uses his notebook to store and retrieve his beliefs. In such cases, Clark argues that Otto's interaction with the notebook enables us to conceive of the notebook as an external memory working *in a way analogous in its effects* to the way many other human beings use their biological memory. The example motivates the introduction of the *parity principle* stating that 'If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognising as part of the cognitive process, then that part of the world is [...] part of the cognitive process' (Clark & Chalmers, 1998). Otto's notebook becomes a vehicle of his mental state because it contains belief-enabling information analogous to neural states in people without Alzheimer's and thereby fulfils the cognitive role of dispositional beliefs. Notice that even if the notebook does not afford the exact same principles of interaction between the cognizer and the database as biological memory and has different physical properties, its function is analogous and – just like biological memory – available for guiding tasks in the world, easily accessible, and automatically endorsed by Otto. Emphasising this functional aspect of the analogy thus seems to meet some of the immediate concerns for drawing a parity between brain-based memory storage and the notebook: just like a notebook can be lost or doubted, we can suffer from forgetfulness, double-check our biological memories employing a calendar and so on (Gallagher, 2011; Sutton et al., 2011; Tollefsen, 2006).

2.2. Extending the individual mind via language

Applying the hypothesis of cognitive extension to the case of language, A. Clark argues that rather than serving merely as a vehicle of already existing symbolic thought, language comes to actually constitute part of the process of thinking (Clark, 1997). It does so not only by virtue of the content of words and sentences, but also crucially by virtue of the *bare materiality* of those words and sentences (Clark, 2006a). The material aspects of language, such as their perceptible depictions or phonemic properties, complement biological processes of cognition by creating new fulcrums of attention, memory and control. A written note enables us to remember a long shopping list more accurately than biological memory alone. A new recipe enables us to finely coordinate our attention and behaviour for the accomplishment of a novel task. Memory and conceptual structures are offloaded in an external linguistic structure. In other words, linguistic patterns enable the cognitive agent to construct, rely upon and manipulate 'cognitive

niches': regularities, affordances and constraints that shape and support cognitive processes thereby functioning as the external vehicles of cognition (Clark, 2006b). However, Clark's proposal of considering language as a case of extended cognition primarily in terms of the materiality of language tends to overshadow the *processes* which make the material aspects of language count as external vehicles of cognition (Steffensen, 2009; Wilson, 2010): the shopping list constitutes an external memory in virtue of enabling and being part of a larger memory process involving writing on a piece of paper and engaging the persisting perceptual cues this provides.

By characterising language as a 'persisting, though never stationary symbolic edifice' (Clark, 2008), Clark focuses on the code-like products of linguistic activities rather than the processes that produce, engage and evolve linguistic forms (Rączaszek-Leonardi & Kelso, 2008; Steffensen, 2009). This focus echoes a more general bias in philosophy of mind, linguistics and cognitive science that implicitly and unreflectively have treated literacy and written text as the prototypical case of language (Linell, 2005). Not neglecting the huge impact of literacy on human culture and cognition (Donald, 2001; Ong, 1982), it is important to acknowledge how written text occupies a rather peripheral and specialised domain of human linguistic behaviour, both in an phylogenetic, ontogenetic and every-day usage-based perspective (Fusaroli & Tylén, 2012; Fusaroli, Rączaszek, & Tylén, 2013). Even when language is written – and under certain respects objectified as autonomously existing material symbols – it is still primarily in support of wider communicative and cognitive practices, that is, to enable interpersonal interaction and coordination. Handbooks are used in teaching and learning practices, holy texts in religious practices, as well as everyday and even legal behaviour, newspapers in the coordination and reshaping of public opinion and so forth. In this sense, language (even when written) is first and foremost a dialogical and intersubjective activity. Language is an *activity* that allows us to coordinate actions, perceptions and attitudes, share experiences and plans, and to construct and maintain complex social relations on different time scales (Kravchenko, 2004, 2007; Rączaszek-Leonardi, 2010; Rączaszek-Leonardi & Cowley, 2012; Thibault, 2011; Tylén et al., 2010).

Let us proceed by degrees towards a description of language as a case of active externalism. Consider first language as skilful coordination of individual behaviour. Notwithstanding his focus on the "symbolic edifice", when Clark concretely accounts for the cognitive benefits of language (or proto-language) he persuasively describes language use as a creative activity achieving cognitive feats impossible to achieve without language (Clark, 2006a). For example, by learning how to employ plastic tokens to label relations of 'sameness' and 'difference', chimpanzees become able to discriminate between, and maybe even conceptualise, relations between relations (Thompson, Oden, & Boysen, 1997). Likewise, by learning to construct

and recall sentences children become capable of combining and remembering several perceptual cues at once (Hermer-Vazquez, Spelke, & Katsnelson, 1999). It would of course be naïve to think that just by possessing a material edge, symbols make these otherwise unachievable feats possible: the plastic token on its own does not enable chimps to conceive meta-relations. Plastic tokens are not *in themselves* cognitive. They are described as cognitive in virtue of their role in the coordination of cognitive activities. Chimps learn – quite laboriously and relying on the reinforcements and the complex social setting of the ethological experiment (cf. Hutchins, 2008) – how to closely coordinate their biological resources with the material features of the plastic tokens so to discriminate between relations and – later on – between meta-relations. Similarly children learn in socially ostensive, and highly scaffolded contexts how to construct sentences (Tomasello, 2003) in order to label and combine multiple perceptual features as well as how to repeat them in order to structure their attention during a task. The real-time feedback loops between the verbal components of language and embodied routines prompts the intuition that it would be unnecessary and unduly complicated to conceive the cognitively potent material vehicles of language as quintessentially external objects, separate from the embodiment of the cognitive agent. As an example, Sutton and co-authors eloquently express the dynamics of language and embodied skilful coping with the world in the case of the experienced cricketer:

'It's true that 'watch the ball' is not an instruction sent from a detached mind to an obedient body, the top-down (re-)programming of the body-machine. The function of the verbal maxim is not exhausted – perhaps even no longer significantly affected – by its semantic content: rather it operates in real time as a *material* symbol, an iterated and interactive self-stimulatory loop. The role of "instructional nudges" like 'watch the ball' or 'jazz hands', as Wheeler suggested, need not be precise control of the microstructure of action: yet the expert performer is using these verbal components of multimodal embodied routines to distribute intelligence, coordinating or often re-setting and rechunking patterns of movement or affect or mood, as one among many forms of scaffolding that support the embodied rebuilding of action sequences from the inside.' (Sutton et al., 2011).

Only as skills are gradually acquired and bodily entrenched do material symbols come to enable and become part of cognitive activities. Or, to put it more provocatively: it is the gradually emerging activity that defines its components as cognitive, with little or no regard to the boundaries of the skull. Language can thus be considered as a particular kind of engaged activity, constituted by practice and skilful coordination, and not just as a set of variously organised material symbols, or even the mere storage or a transfer of information via material symbols.

While the notion of *language as a skilful activity*, that is, something we do, has been somewhat overshadowed by the emphasis on materiality, it still has found some articulation in the extended mind hypothesis. However, a second crucial aspect has been completely neglected, though it complements and productively develops the mode: *language as something we do together*, as an intersubjective dialogical activity (Fusaroli, Demuru, & Borghi, 2012).

3. From material symbols to dialogical minds

3.1. Language as a tool for interacting minds

Focusing on linguistic activities, we notice how Clark's solitary individuals employing linguistic actions – repeating instructions, taking notes, talking to themselves, etc. – seem rather the special case than a representative example of linguistically mediated cognitive processes. A clue to the core cognitive nature of language is to be found in the way we evolve, learn to produce and interpret material symbols in the first place. A growing number of studies (Donald, 1991; Galantucci, 2009; Loreto & Steels, 2007; Tomasello et al., 2005) are indeed suggesting that language evolved (and keeps evolving) from a pressure for increasingly sophisticated means of socio-cultural coordination and cooperation: by producing material symbols, humans manipulate public cognitive niches open – at least potentially – to other individuals thereby enabling intersubjective cognitive activities (Fusaroli & Tylén, 2012; Maturana & Varela, 1980; Fusaroli et al., 2013). Examples of material symbols as the vehicles of intersubjective cognitive coordination are pervasive, ranging from everyday chats and gossip to ancient manuscripts enabling the coordination of knowledge and activities across millennia. However, the emphasis on the socially coordinative role of language brings about an important shift in the language-as-tool metaphor compared to Clark's account: it is not so much a question of coupling between an individual and the material symbols of language, but rather how language as a dynamic activity enables individuals to form intersubjective cognitive systems. In other words, the cognitive extension is thus not the materiality of language *per se*, but the way this materiality enables the dialogical engagement with an interlocutor.

We are not the first to suggest that the extended mind hypothesis applies to interpersonal interactions. Tollefsen (2006) and Theiner et al. (2010) have argued for *collective minds*, relying on the case of *transactive memory*: long-term interpersonal relationships can lead to a distribution of information storage and retrieval across partners: e.g. one partner remembering time and date of events, the other how to drive to their location, each of them contributing to mark things as to be remembered and cueing each other in remembering them correctly. Transactive memory is functionally equivalent to biological memory and to the use of a notebook as memory support: the partners are mostly available and accessible to each other and they

actively create, maintain and retrieve endorsed information in a way that makes them complementary (Peltokorpi, 2008; Sutton et al., 2011; Theiner et al., 2010). Along similar lines, Tollefsen and Dale discuss how low-level behavioural coordination creates in certain cases a feeling of joint agency between the individuals involved (Tollefsen & Dale, 2012).

By arguing that language can give rise to dialogically extended minds, we not only intend to contribute to current models of cognitive extension by adding to the range of phenomena describable as collective minds, but also we point to some of the interaction processes through which collective minds are created and maintained.

3.2. Empirical support for dialogically extended minds

With a blend of conceptual and empirical arguments, Tollefsen and Dale (2012) make a compelling case that joint agency is related to basic coordinative mechanisms. Indeed, individuals have been argued to be strongly other-oriented: the mere presence of another person has been shown to have motivational effects (Brennan, Galati, & Kuhlen, 2010). During interpersonal interactions, people readily adapt to each other by continuous reciprocal priming and mimicking (Chartrand & Bargh, 1999; Pickering & Ferreira, 2008). Via progressive behavioural alignment, interacting individuals often feel increased emotional attachment (Marsh, Richardson, & Schmidt, 2009) and begin to share higher cognitive processes (Hasson et al., 2012). Even in absence of rich behavioural interaction, people have been argued to be very sensitive to the other's perspectives and engagement (De Jaegher, Di Paolo, & Gallagher, 2010; Di Paolo & De Jaegher, 2012; Duran & Dale, 2013; Tylén et al., 2013). Similar observations have been made with respect to linguistic dialogue. Here the linguistic aspects of language are considered in continuation with other expressive behaviours such as gestures, postures and facial expressions. The term 'language' is thus used quite broadly about the simultaneous orchestration of many modalities all of which can be considered public, material instantiations of the conversation as a coordinative activity (Goodwin, 2011). Conversational analysts have shown the complexity of conversational moves and how conversations rely on functionally structured sequences of speech turns (Clark, 2005; Schegloff, 1986, 2004). Individuals' behaviour in conversation is not free-floating, but typically fulfil roles in larger *interactional scripts*, which act as scaffolding and constraining the possibilities of actions and interpretation in joint activities (Clark, 1996; Levinson, 1983; Mills, 2013; Rączaszek-Leonardi & Cowley, 2012; Thibault, 2011). By relying on such structures, individuals who suffered brain damage have been shown able to coordinate with their interlocutors in order to accomplish otherwise unfeasible cognitive and communicational tasks (Goodwin, 2003, 2011). Language has thus been argued to be an invaluable tool for interacting minds (Tylén et al., 2010).

Building on these ideas, the growing field of experimental pragmatics and semiotics (Galantucci & Garrod, 2010) persuasively shows how humans develop and maintain signs and language as a function of their coordinative needs. In other words, language is continuously carved as a social coordinative device (Fusaroli & Tylén, 2012), which enables people to do things better together, or even to do things they would not be able to do as individuals. Using language, pairs of interlocutors manage to easily navigate otherwise unsolvable mazes (Galantucci, 2005; Garrod & Anderson, 1987; Mills, 2011); to agree on references without being co-present and sharing the same visual scene (Dale, Kirkham, & Richardson, 2011; Shockley, Richardson, & Dale, 2009) and to co-construct shared conceptual schemes for spatial navigation (Anderson et al., 1991; Fusaroli et al., 2012; Garrod & Doherty, 1994). The crucial thing in these studies is that they enable us to go beyond the simple idea that language is used to coordinate an individual's activities to the study of how linguistic intersubjective engagement generates interpersonal synergies. That is, how do the actions and cognitive processes of interlocutors come to constitute a single coordinated system able to do better or more than the individual components (Dale et al., 2013; Latash, Scholz, & Schoner, 2007; Riley et al., 2011; Fusaroli et al., 2013)? Conversations have been observed to have intersubjective coordinative effects on a wide variety of behavioural measures. For example, Shockley, Santana, and Fowler (2003) and Shockley (2007) set up a puzzle task in which the members of a pair discussed an array of pictures to jointly identify the differences. The authors found that during conversation the participants' postural sway trajectories – the subtle, continuous motion of the body during upright standing – became more and more coordinated over time, even when they could not see each other. Analogously, in a study where interlocutors were asked to discuss pictures of characters from popular TV shows and surrealist paintings (Richardson & Dale, 2005; Richardson, Dale, & Kirkham, 2007), their eye movements became strongly coordinated, so that interlocutors were highly likely to look at the same thing at the same point in time. This coordination of gaze is found even if, after an initial exposure to the pictures, only one of the interlocutors keeps having visual access to it (Richardson, Dale, & Tomlinson, 2009), which suggests that the effect is not driven solely by properties of the environmental stimuli, but in important ways by the linguistic engagement.

Crucially, however, language does more than simply enabling the constitution of behavioural synergies: it also enables the constitution of *informational synergies*, or what we term *dialogically extended minds*. A clear example of this can be found in a recent study on collective decision-making (Bahrami et al., 2010). Bahrami and colleagues were inspired by studies on how experimental participants would enhance their task performance when – rather than relying on a single sensory modality – they could integrate information from different sensory modalities (Ernst &

Banks, 2002). In these cases information were internally combined within the same biological brain. However, the researchers wanted to explore the possibilities for integration of sensory information *between* individuals. In this study linguistic dialogue constituted the 'coupling link' between two individual sensory systems.

In the experiment, pairs of participants were instructed to individually indicate in which of two brief visual displays they had just been shown a contrast oddball. When their individual decisions diverged, they were prompted to discuss and report a joint decision. In order for a pair to achieve a cooperative benefit, that is, to perform better than the best of the two individuals, they had to find ways of assessing and comparing their individual levels of confidence so as to choose – on a trial-by-trial basis – the decision of the more confident participant. The results showed that on average pairs succeeded in doing so. In other words, by interacting linguistically the participants were able to effectively share and combine information, making collective decisions more effective. In analogy to Ernst and Banks (2002), linguistic dialogue enabled the two participants to act like one intersubjective oddball-identifying system. However, not all pairs performed equally well and actually reached a benefit from their cooperation. This points to the issue of *skilfulness* in coordination. In a follow-up study, Fusaroli et al. (2012) investigated which properties of the dialogical dynamics could be associated with success in the task and observed substantial variability in pairs' abilities to coordinate linguistically.

Consider the following excerpt, focusing on the expressions of confidence:

Danish original transcription	English translation
B: ((laughs)) jeg ved det ikke	B: ((laughs)) I don't know
A: jeg ved det heller ikke. Jeg så både i venstre hjørne og midt for til højre på dem begge	A: I don't know either. I saw something both in the left corner and in the centre on the right in both of them
B: okay jeg synes det var ovre i venstre side men uhm pas	B: okay, I think it was over in the left side but uhm I'll pass

The participants generally use a variety of everyday expressions such as 'I don't know', 'I saw something' or 'I think it was ...' when talking about their levels of confidence. It is also possible to notice that when interlocutor B employs 'I don't know', A's successive expression of confidence also employs 'to know'. This phenomenon, dubbed "local linguistic alignment" was quite widespread: participant pairs generally tended to adapt to each other's way of talking about confidence on a trial-by-trial basis. By this kind of mutual alignment, the participants got to share a common language in which to express and compare confidence: a quite crucial condition for creating an informational synergy in this task. But they did so at different degrees, some pairs showing higher and other lower transi-

tion probabilities, that is, probabilities that a given confidence expression would be repeated by the other participant in the next interaction. This measure was shown to significantly correlate with how well the pairs would perform in the joint decision task. Indeed, the more they skilfully adapted their confidence expressions to each other, the higher benefit did the pairs achieve from their cooperation.

This is a clear example of how language can be conceived of as a process of contextually sensitive reciprocal adaptation, to the point that the dynamics of the engagement evolve and shape words and other expressive behaviours (Tylén et al., 2013). While the perceivable ‘material’ aspects of language arguably constitute the public arena for collective cognition in this experiment (corresponding to Clark’s ideas), it is not so much the material symbols as such that do the trick, but the ongoing intersubjective engagement that material symbols make possible. Two further findings can be brought to strengthen this point. Different pairs evolved and employed very different sets of expressions to talk about confidence. However, the different particular lexicalizations of confidence did not seem to make a difference. Rather, it was the degree to which the participants flexibly employed and reciprocally adapted their confidence expressions that correlated with task performance. In other words, it is not language as particular material symbols that constitutes the basic cognitive extension in this case; rather, it is through the *intersubjective interaction* and *coordination* of the material symbols that individual participants become each other’s cognitive extensions. Indeed, a recent variation of the experiment further supports this claim. Here, the experimenters introduced a condition where the participants’ possibilities for communication were restricted to a pre-given set of material symbols: a numerical scale from 1 (indicating ‘very doubtful’) to 5 (indicating ‘absolutely sure’), through which the individuals had to indicate their confidence (Bahrami et al., 2012). While material symbols were undoubtedly available, the possibilities for reciprocal adaptivity and engagement were strongly reduced. Interestingly, in these conditions, the pairs failed to achieve the same amount of efficacy in integrating information, displaying significantly lower cooperative benefit in their joint performance. This seems to further articulate the point that language should not be seen as simply a conventional code for transmitting information, nor a set of ready-to-go cognition-enhancing material symbols. More crucially, the cognitive enhancing potential of language lies in the skilful intersubjective engagement with public, expressive resources that facilitates informational-synergy-creating intersubjective coupling.

One last study (Dale et al., 2011) brings informational and behavioural synergies together. It employs the ‘tangram task’ (Krauss & Weinheimer, 1964) in which two participants are presented with the same set of abstract shapes portrayed in different orders, while their eye-movements are recorded. One participant instructs the other to arrange

his shapes so that the orders match. To accomplish this task, the participants must find a way to refer to the abstract shapes. Like in the visual discrimination task above, the participants adapt to each other and stabilize jointly constructed descriptions for the shapes (Clark & Wilkes-Gibbs, 1986). As this happens, participants take less time to solve the task and require fewer words to do so. By developing, sharing, stabilizing and engaging linguistic forms to refer to the shapes, the interlocutors hugely increase their ability to solve the task and the coordination between their behaviours. Indeed, in the beginning of the tangram task, when director and matcher have not yet become linguistically well coordinated, the eyes of the matching participant follow the eyes of the directing participant with a distinctive delay in time. However, as they jointly construct shared referential expressions, participants’ visual attention become increasingly coordinated: directing and matching participants look at the same things at the same time, indicating that they sample the world in increasingly similar and effective ways. In other words, they become one coherent, functionally coupled cognitive unit able to accomplish feats the individuals alone could not, a dialogically extended mind.

In both the visual discrimination and tangram examples through repeated interactions, interacting agents gradually develop stable linguistic (or symbolic) structures that allow them to achieve new epistemic ends. The entrenchment and skilful engagement of these emergent repertoires of shared symbols enables individuals to engage in a plurality of otherwise impossible activities: the joint construction and active manipulation of cognitive niches (Clark, 2006b; Sinha, 2009). By constructing and engaging graded expressions of confidence or labels for abstract shapes, interlocutors create a set of constraints on which they can rely in their joint cognitive activity, a cognitive niche that would be difficult to imagine without the aid of language. In these cases, and many others cf. e.g. interesting works on the cognitive functions of narrative, (Donald, 2001), language structures shared situation models, distributes roles in the on-going activity, and guides and constrains the possibilities for coordinated attention, action and cognition. It allows the interlocutors to gradually align on stabilised but flexible shared leads and constraints for intersubjective coordination. Interestingly, long-term relationships are not necessary to form interpersonal cognitive systems: dialogical exchanges can in short time enable the efficient co-construction of information and procedures to accomplish more efficient or innovative cognitive tasks.

3.3. Dialogical extended minds: discussion and potential critiques

In the previous paragraphs we have argued that interlocutors skilfully engaged in dialogue tend to form behavioural and informational synergies. On the one hand, these work in ways analogous to individual cognitive systems, but on the other hand, they can even improve the

cognitive performance beyond their constituent components (the individual interlocutors). Our main purpose thus was to demonstrate how Clark's proposal of language as a case of cognitive extension is further strengthened by focusing on the social aspects of language. Moreover, we draw on the concept of collective minds, the conceptual soundness of which rests on promising grounds (Carr, 1986; Fusaroli, Granelli, & Paolucci, 2011; Gallagher, 2011; Gallagher & Crisafi, 2009; Theiner et al., 2010; Tollefsen, 2006; Zahavi, 2007). Of course, these arguments could be subjected to all the usual critiques the extended mind hypothesis has met during the last couple of decades. While it is beyond the scope of this article to engage the whole debate on "the mark of the mental" (Adams & Aizawa, 2010; Rupert, 2011), we will focus on two prominent objections, motivated by our emphasis on "interpersonal linguistic engagement" and by our willingness to accept on-the-fly linguistic interpersonal synergies: the "causal coupling-constitution fallacy" (Block, 2005) and the "cognitive bloat" argument (Adams & Aizawa, 2008; Rupert, 2009).

The first critique states that causal coupling is not constitution, in other words, that the causal connection of two systems is not enough to warrant the claim that they constitute *one* cognitive system. Our argument is not affected by this critique, since we are not advocating any simple causal coupling of two systems. Our proposal describes the interaction between two cognitive agents and the enabling of new informational synergies by the interaction. The reciprocal engagement involved in interpersonal synergies does not simply create a one-way causal connection, but – more interestingly – a reciprocal interdependence between the two systems. This interdependence constrains the interlocutors individual possibilities for action (e.g. by aligning their attention, or how they evaluate their confidence) and at the same time forms higher-order cognitive units with improved or innovative capabilities as compared to its components (Fusaroli & Tylén, submitted for publication; Fusaroli et al., 2013; Kelso, 2009; Riley et al., 2011; Fusaroli et al., 2013).

The second possible critique is the "cognitive bloat" argument: by allowing the extension of the term "cognitive" beyond the boundaries of skull and skin, we create a category that encompasses everything and therefore is explanatorily useless (Adams & Aizawa, 2008; Rupert, 2009). We argue that this is not the case. On the contrary, not all conversational arrays automatically come to constitute interpersonal synergies. This requires a certain level of skilful linguistic engagement. As evident from the empirical studies reviewed, synergy effects can be achieved to a lower or higher degree. In other words, interlocutors can become more or less close to an ideal model of dialogical mind according to the dynamics at play. Importantly, we also suggest a possible mechanism for the creation and maintenance of dialogical minds: the co-construction of interactional routines, such as the context-sensitive alignment of expressive behaviours. The

identification of such mechanisms enabling interpersonal synergies motivates new experimentally testable hypotheses (Dale et al., 2013). For instance, we predict that interlocutors in competition or conflict display a significantly lower alignment than cooperating interlocutors. Preliminary studies by Paxton and Dale seem to confirm the prediction (Paxton & Dale, submitted for publication). Accordingly, we argue that the study of collective minds can and indeed should be articulated with the empirical findings from the study of joint action (Galantucci, 2009; Hasson et al., 2012; Obhi & Sebanz, 2011; Ramenzoni et al., 2012; Shockley et al., 2009; Tollefsen & Dale, 2012). Tollefsen and Dale (2012) have sketched how non-verbal alignment could be connected with joint agency. We have articulated how linguistic coordination possibly relates to collective action and cognition. Obviously, these are only initial studies. More empirical and conceptual research is needed in order to better understand the mechanisms involved in the creation and maintenance of collective minds, such as emotional affiliation (Marsh et al., 2009), perspective-taking (Brennan et al., 2010; Duran & Dale, in press), alignment (Fusaroli & Tylén, 2012) and interactional routines (Fusaroli et al., 2013). Finally, our proposal potentially points to an even more pronounced role for social world in enabling an individual's core cognitive capacities. This is an interesting venue for future empirical research.

4. Conclusions

In line with A. Clark, we consider language to be a prominent case of extended cognition. However, we claim that the true power of language cannot be grasped by focussing primarily on its material properties, nor on how they facilitate individual reasoning. Language is first and foremost intersubjective engagement. Its material and symbolic aspects thus constitute the public arena for dynamical, interpersonal synergies. Language enables individuals to coordinate their cognitive processes in evolutionarily unprecedented ways, effectively constituting dialogically extended minds. In the skilful intersubjective engagement of symbolic patterns, human beings rely on each other and on established cultural practices to achieve cognitive feats that would otherwise be beyond reach.

Acknowledging these intersubjective and cultural dimensions of the way language constitutes interpersonal cognitive systems, new lines of interdisciplinary research are opened in the field of extended and social cognition. In this paper we focused on language as a cognitive phenomenon that combines core aspects of intersubjectivity with the creation and maintenance of external vehicles of cognition. In this perspective, the extended mind hypothesis gets to be supported and further articulated by on-going research in social cognition investigating the constitution and dynamics of embodied intersubjective engagement (Gangopadhyay, 2011). At the same time, research in social cognition is complemented by a focus on how dialogical

engagement comes to constitute extended cognitive processes. Last but not least, our perspective fills a gap in the present debate by providing an initial account of the pervasive role of language in creating intersubjective coupling and thereby shared cognitive processes.

Acknowledgements

Research was funded by the Danish Council for Independent Research - Humanities (FKK) project “Joint Diagrammatical Reasoning in Language”, and the EUROCORES grant EuroUnderstanding “Digging for the Roots of Understanding”.

References

- Adams, F., & Aizawa, K. (2008). *The bounds of cognition* (vol. xiii). Malden, MA: Blackwell Pub., 197p.
- Adams, F., & Aizawa, K. (2010). Defending the bounds of cognition. In R. Menary (Ed.), *The extended mind* (pp. 67–80). Cambridge, Mass: MIT Press.
- Anderson, A. H. et al. (1991). The HCRC map task corpus. *Language and Speech*, 34(4), 351.
- Bahrami, B. et al. (2010). Optimally interacting minds. *Science*, 329(5995), 1081–1085.
- Bahrami, B. et al. (2012). What failure in collective decision-making tells us about metacognition. *Philosophical Transactions of the Royal Society B Biological Sciences*, 367(1594), 1350–1365.
- Block, N. (2005). Review of Alva Noe, Action in perception. *Journal of Philosophy*, 102, 259–272.
- Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, 53, 301–344.
- Carr, D. (1986). Cogitamus Ergo Sumus: The intentionality of the first-person plural. *Monist*, 69(4), 523.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893–910.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, A. (1997). *Being there: Putting brain body and world together again* (vol. xix). Cambridge, Mass: MIT Press, 269p [4] p. of plates.
- Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, 7(4–5), 507–525.
- Clark, A. (2006a). Material symbols. *Philosophical Psychology*, 19(3), 291–307.
- Clark, A. (2006b). Language, embodiment, and the cognitive niche. *Trends in Cognitive Sciences*, 10(8), 370–374.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. *Philosophy of Mind* (vol. xxix). Oxford; New York: Oxford University Press, 286p.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39.
- Dale, R. et al. (2013). The self-organization of human interaction. *Psychology of Learning and Motivation*, 59, 43–95.
- Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in Psychology*, 2, 355.
- De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 441–447.
- Di Paolo, E., & De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in Human Neuroscience*, 6, 163.
- Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition* (vol. viii). Cambridge, Mass: Harvard University Press, 413p.
- Donald, M. (2001). *A mind so rare: The evolution of human consciousness* (vol. xiv, 1st ed.). New York: Norton, 371p.
- Duran, N. D., & Dale, R. (2013). Perspective-taking in dialogue as self-organization under social constraints. *New Ideas in Psychology*, xx(xx), xxx–xxx.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433.
- Fodor, J. A. (1975). *The language of thought. The language & thought series* (vol. x). New York: Crowell, 214p.
- Fodor, J. A. (2008). LOT 2: The language of thought revisited. Oxford & New York: Oxford University Press.
- Fusaroli, R. (2011). The social horizon of embodied language and material symbols. *Versus*, 112–113, 95–120.
- Fusaroli, R. et al. (2012). Coming to terms: An experimental quantification of the coordinative benefits of linguistic interaction. *Psychological Science*, 23(8).
- Fusaroli, R., & Tylén, K. (submitted for publication). *Individual behavior, interactive alignment or interpersonal synergy? A model-comparison study on linguistic dialog*.
- Fusaroli, R., Demuru, P., & Borghi, A. M. (2012). The intersubjectivity of embodiment. *Journal of Cognitive Semiotics*, 4(1).
- Fusaroli, R., et al. (2013). *Conversation, coupling and complexity: Matching scaling laws predict performance in a joint decision task*, Poster presented at CogSci 2013, Berlin.
- Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2013). Dialogue as interpersonal synergy. *New Ideas in Psychology*, in press.
- Fusaroli, R., Granelli, T., & Paolucci, C. (2011). *The external mind (thematic issue of Versus)* (vol. 112–113). Milano: Bompiani.
- Fusaroli, R., & Tylén, K. (2012). Carving language for social interaction: A dynamic approach. *Interaction Studies*, 13(1), 103–123.
- Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29, 737–767.
- Galantucci, B. (2009). Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, 1(2), 393–410.
- Galantucci, B., & Garrod, S. (2010). Experimental semiotics: A new approach for studying the emergence and the evolution of human communication. *Interaction Studies*, 11(1), 1–13.
- Gallagher, S. (2011). The overextended mind. *Versus*, 112–113, 57–68.
- Gallagher, S., & Crisafi, A. (2009). Mental institutions. *Topoi – An International Review of Philosophy*, 28(1), 45–51.
- Gangopadhyay, N. (2011). The extended mind: Born to be wild? A lesson from action-understanding. *Phenomenology and the Cognitive Sciences*, 10, 377–397.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27(2), 181–218.
- Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, 53(3), 181–215.
- Geeraerts, D., & Cuyckens, H. (2007). *The Oxford handbook of cognitive linguistics*. Oxford handbooks, p. 1334.
- Goodwin, C. (2003). *Conversation and brain damage* (vol. x). Oxford; New York: Oxford University Press, 314p.
- Goodwin, C. (2011). Building action in public environments with diverse semiotic resources. *Versus*, 112–113.
- Habermas, J., & Cooke, M. (1998). *On the pragmatics of communication. Studies in contemporary German social thought* (vol. viii). Cambridge, Mass: MIT Press, 454p.
- Hasson, U. et al. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*.
- Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology*, 39(1), 3–36.

- Hutchins, E. (2008). The role of cultural practices in the emergence of modern human intelligence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1499), 2011–2019.
- Kelso, J. A. S. (2009). Synergies: Atoms of brain and behavior. In D. Sternad (Ed.), *Progress in motor control: A multidisciplinary perspective* (pp. 83–91). New York: Springer.
- Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction. *Psychonomic Science*, 113–114.
- Kravchenko, A. V. (2007). Essential properties of language, or, why language is not a code. *Language Sciences*, 29(5), 650–671.
- Latash, M. L., Scholz, J. P., & Schoner, G. (2007). Toward a new theory of motor synergies. *Motor Control*, 11(3), 276–308.
- Levinson, S. C. (1983). *Pragmatics. Cambridge textbooks in linguistics* (vol. xvi). Cambridge: Cambridge University Press, 420p.
- Linell, P. (2005). The written language bias in linguistics: Its nature, origins and transformations. New York: Routledge.
- Loreto, V., & Steels, L. (2007). Social dynamics – Emergence of language. *Nature Physics*, 3(11), 758–760.
- Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, 1(2), 320–339.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living* (vol. 42). Springer.
- Mills, G. (2011). The emergence of procedural conventions in dialogue. In 33rd Annual conference of the cognitive science society. Boston, USA.
- Mills, G. (2013). Dialogue in joint activity: Complementarity, convergence and conventionalization. *New Ideas in Psychology*, xxx(xx), xxx–xxx.
- Obhi, S. S., & Sebanz, N. (2011). Moving together: Toward understanding the mechanisms of joint action. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, 211(3–4), 329–336.
- Ong, W. J. (1982). *Orality and literacy: The technologizing of the word. New accents* (vol. x). London; New York: Methuen, 201p.
- Paxton, A., & Dale, R. (submitted for publication). *Argument disrupts interpersonal alignment*.
- Peltokorpi, V. (2008). Transactive memory systems. *Review of General Psychology*, 12(4), 378–394.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3), 427–459.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2), 169–190, Discussion 190–226.
- Rączaszek-Leonardi, J. (2010). Multiple time-scales of language dynamics: An example from psycholinguistics. *Ecological Psychology*, 22(4), 269–285.
- Rączaszek-Leonardi, J., & Cowley, S. J. (2012). The evolution of language as controlled collectivity. *Interaction Studies*, 13(1), 1–16.
- Rączaszek-Leonardi, J., & Kelso, J. A. S. (2008). Reconciling symbolic and dynamic aspects of language: Toward a dynamic psycholinguistics. *New Ideas in Psychology*, 26(2), 193–207.
- Ramenzoni, V. C. et al. (2012). Interpersonal and intrapersonal coordinative modes for joint and single task performance. *Human Movement Science*.
- Richardson, D. C., & 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 Science*, 18(5), 407–413.
- Richardson, D. C., Dale, R., & Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science*, 33(8), 1468–1482.
- Riley, M. A. et al. (2011). Interpersonal synergies. *Frontiers in Psychology*, 2, 38.
- Rupert, R. D. (2009). *Cognitive systems and the extended mind. Philosophy of mind* (vol. xiii). Oxford; New York: Oxford University Press, 268p.
- Rupert, R. D. (2011). Empirical arguments for group minds: A critical appraisal. *Philosophy Compass*, 6(9), 630–639.
- Schegloff, E. A. (1986). The routine as achievement. *Human Studies*, 9(2–3), 111–151.
- Schegloff, E. A. (2004). Putting the interaction back into dialogue. *Behavioral and Brain Sciences*, 27(2), 207–208.
- Shockley, K. et al. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology. Human Perception and Performance*, 33(1), 201–208.
- Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, 1(2), 305–319.
- 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.
- Sinha, C. (2009). Language as biocultural niche and social institution. In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics*. Amsterdam; Philadelphia: John Benjamins Pub. Co..
- Steffensen, S. V. (2009). Language, languaging, and the extended mind hypothesis. *Pragmatics & Cognition*, 17(3), 677–697.
- Sutton, J. et al. (2011). The psychology of memory, extended cognition and socially distributed remembering. *Phenomenology and the Cognitive Sciences*, 9(4), 521–560.
- Sutton, J. et al. (2011). Applying intelligence to the reflexes. *Journal of the British Society for Phenomenology*, 42(1), 78–103.
- Theiner, G., Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research*, 11(4), 378–395.
- Thibault, P. J. (2011). First-order languaging dynamics and second-order language: The distributed language view. *Ecological Psychology*, 23(3), 210–245.
- Thompson, R., Oden, D., & Boysen, S. (1997). Language-naïve chimpanzees (Pan troglodytes) judge relations between relations in conceptual matching-to-sample task. *Journal of Experimental Psychology – Animal Behavior Processes*, 23(1), 31–43.
- Tollefsen, D. P. (2006). From extended mind to collective mind. *Cognitive Systems Research*, 7(2–3), 140–150.
- Tollefsen, D., & Dale, R. (2012). Naturalizing joint action: A process-based approach. *Philosophical Psychology*, 25(3), 385–407.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition* (vol. viii). Cambridge, Mass: Harvard University Press, 388p.
- Tomasello, M. et al. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675–691, discussion 691–735.
- Tylén, K. et al. (2010). Language as a tool for interacting minds. *Mind & Language*, 25(1), 3–29.
- Tylén, K. et al. (2013). Making sense together: A dynamical account of linguistic meaning making. *Semiotica*, 194, 39–62.
- Tylén, K. (2012). Interaction vs. observation: distinctive modes of social cognition in human brain and behavior? A combined fMRI and eye-tracking study. *Front. Hum. Neurosci.*, 6, 331.
- Vygotsky, L. S. (1978). *Mind and society*. Cambridge, MA: Harvard University Press.
- Wilson, R. A. (2010). Meaning making and the mind of the externalist. In R. Menary (Ed.), *The extended mind* (pp. 167–188). MIT Press.
- Donald, M. (2001). *A mind so rare: The evolution of human consciousness*.
- Kravchenko, A. V. (2004). *Essential properties of language from the point of view of autopoiesis*.
- Zahavi, D. (2007). Subjectivity and the first-person perspective. *The Southern Journal of Philosophy*, 45, 66–84.