Chapter 2
Human Agency and the Resources of Reason

Martin Neumann and Stephen J. Cowley

Abstract The evidence shows that human primates become (relatively) rational actors. Using a distributed perspective, we identify aspects of human agency that raise important questions for sociocognitive science. In humans, we argue, agency does not centre on individual agents. Cognitive, social and linguistic outcomes depend on skills in moving in and out of aggregates that bind people, artifacts, language and institutions. While recognising the value of symbol processing models, these presuppose the embodied events of *human symbol grounding*. At a micro level, humans coordinate with others and the world to *self-construct* by cognising, talking and orienting to social affordances. We trace the necessary skills to sense-saturated coordination or *interactivity*. As a result of perceiving and acting on the environment, human individuals use the artificial to extend their natural powers. By using verbal patterns, artifacts and institutions, we become imperfect rational actors whose lives span the micro and the macro worlds.

Towards Sociocognitive Science

Humans make extensive use of cultural resources that include languages. Alone of the primates, their infants develop in a normative world that allows species to develop a distinctive type of agency. This is manifest individually and, just as strikingly, in how human groups implement long-term plans. For historical reasons, attempts to explain human agency have typically sought its origins in, not a history of

M. Neumann
University of Koblenz, Koblenz, Germany
e-mail: maneumann@uni-koblenz.de

S.J. Cowley (✉)
School of Psychology, University of Hertfordshire, College Lane, Hatfield, Hertfordshire,
AL10 9AB, UK
e-mail: s.j.cowley@herts.ac.uk

S.J. Cowley
Department of Language and Communication, University of Southern Denmark,
Sdr. Stationsvej 28, 4200 Slagelse, Denmark

acting in the world, but genes and/or brains. While biological factors are crucial, we argue that their function is to allow living beings to exploit the environment to develop cognitive, social and linguistic skills. The process arises as infants learn to talk or, in exceptional cases, make use of signing. We therefore argue that it is through participating in language-activity that human primates *become* rational actors. Ours is thus a reworking of what makes us human in terms of Durkheim's (1895 [1998]) famous claim that only the social can ground the social.

Embodied coordination enables infants to self-construct as members of a culturally extended ecology. Once we recognise that social predispositions for embodied coordination are functionally reorganized by encultured body-world activity, Durkheim's view ceases to appear circular. Social behaviour arises as we are moved by others to coordinate in increasingly complex ways. By hypothesis, bodies self-organise by learning to cognise, speak and act strategically: human agency develops within aggregates that bind together artifacts, institutions, and ways of experiencing the world. It is a human capacity for moving in and out of such aggregates—for exploiting embodied coordination—that enables social action to drive the self-construction of rational human actors. In tracing human agency to cognition beyond the brain, a history of coordination is seen as the basis of knowledge. Infants use spontaneous activity in learning to orient to a population's linguistic and other social practices. On the distributed perspective, this shapes human agency—infants discover ways of achieving effects that are, inseparably, cognitive, linguistic and sociological.

The Distributed Perspective

Human agency has previously been traced to how acting and thinking are honed by the demands of sociocultural environments. This is done in, for example, activity theory, cultural psychology and the pragmatism of Mead. What is novel to the distributed perspective is the view that agency results from acting and perceiving in socially distributed systems. For readers of *Cognition Beyond the Brain*, the idea will be familiar. The agency of a pilot who lands a plane is non-local in that, as Hutchins (1995) shows, it is distributed across bodies (and brains) that coordinate an aggregate of resources. The pilot uses readings from the instrument panel, messages from ground control, and real time interaction with a co-pilot. Far from centering on a neural system, agency arises in *acting* with material, cultural, and social structures. Ethnography makes that clear. To understand how such aggregates function, however, systemic output must be separated from construals and associated actions. In Tetris, human-computer aggregates rely on both goal-directed actions and epistemic actions that use the world to simplify cognitive tasks (Kirsh and Maglio 1994). Epistemic actions depend on sense-saturated coordination or player-computer *interactivity*. They change on-screen resources in ways that suit the player's expert knowledge. In conversations, tight coupling allows people to concert by using, not just what is said, but also how speech and movements are integrated in rapid or

pico time-scales.[1] In tight coupling that uses cultural resources, in Cronin's (2004) terms, we are cognitively cantilevered into the *Umwelt*. Coordination that is faster than conscious perception drives spontaneity by linking expertise with the results of joint action. The thesis of our paper is that this sense-saturated interactivity shapes human agency by giving us skills in dealing with artifacts, people, and languages.

In ethnography, language is identified with the words that are actually spoken (and which can be transcribed). However, its functionality depends on the fact that language too is *distributed* (see, Cowley 2007c, 2011a; Thibault 2011). In its paradigmatic face-to-face settings, language contributes to action by social aggregates: it is activity in which human beings re-enact the cultural practices and patterns of the ecology. Since it is both situated and verbal, people draw on each other's thinking. This is possible because, unlike animal signalling, languaging has a non-local aspect. Embodied activity links circumstances to past situations by virtue of how we perceive verbal patterns. During languaging, cultural constraints prompt real-time construal. In conversations, verbal patterns arise as concerted acts of articulation (or signing) are accompanied by facial and other expression. Since the results are both embodied and non-local, infants can link metabolic dynamics with between-people coupling. Without hearing what is said, interactivity—and the feeling of thinking—sensitise them to normative resources. While initially reliant on circumstances, they gradually come to hear language as consisting in thing-like parts or to *take a language stance* (Cowley 2011b). Once utterances sound like utterances of something, perceived results or *wordings* can be used in, for example, asking about things. Language, agency and rationality are, irreducibly, individual *and* collective. Cognition links the world in the head with the physical, linguistic and cultural processes of an extended ecology (Steffensen 2011)—a place where individual actions carry cultural value. Human agency self-organises during individual development. Infants use circumstances to become intentional and, later, make use of the resources of reason. Since language is ecological, dialogical and (partly) non-local, rationality co-emerges with individual agency. Although based in interactivity, ways of feeling out the world are supplemented by intelligent—partly conformist—use of external resources.

Agency and Human Agency

The term *agency* can be applied to people, animals, social institutions and inorganic processes. In the first place, it is therefore useful to distinguish physico-chemical agents and living systems. Only in biology do systems (and lineages of systems) set parameters (see, Pattee 1969, 1997) that allow them to measure and control aspects of their environments. Living systems are adaptive and yet also able to maintain autonomy (see, Di Paolo 2005). Adaptive self-organisation allows even single celled

---

[1]This is the time-scale within which gestures are made and syllables articulated. It can broadly be associated with a window of around 200 ms. It is especially important in prosody (see, Cowley 2009) but, at this scale, interactivity is full-bodied (for detailed discussion, see Thibault 2011).

bacteria to explore their worlds by linking genes and metabolism with viable use of available resources. Flexibility increases in living-systems that use brains, development and learning. Yet, these processes too depend on self-organisation or how organic systems exploit the world beyond the body (including other organisms; Thompson 2007). Organisms are aggregated systems whose parameters link a lineage's history with, in embrained species, experiential history of encountering the world.

In evolutionary time, organisms show flexibility as they adapt to the world and, more strikingly, adapt the world to them. For Jarvilehto (1998, 2009), the world of the living is to be conceptualised in terms of interlocking Organism-Environment Systems (O-E systems). The necessity of the view appears with perceptual learning: as Raguso and Willis (2002) show, foraging hawk-moths learn about local sets of flora. In optimising their behaviour, their partly plastic brains link genetically based self-organisation with learning about an environment. Further, as conditions change, they alter the parameters. The example is apposite in that, while such intelligence is organism-centred, this does not apply to all insects of comparable neural dimensions. In bees and other eusocial insects, cognition serves the colony rather than individuals. Below we argue that languages, technology and money make humans partly eusocial. Since we live in a culturally extended ecology, we are *hyper-social* beings (Ross 2007) whose primate intelligence is extended as we orient to eusocial resources that function as cultural (second-order) constraints. Individual-environment relations thus transform individual experience, learning and ontogeny.

Human uniqueness depends on neither our hypersocial nature nor our propensity to exploit the world beyond the brain. What is most striking about human agency is how it combines artificial rigidity with our natural flexibility. As organism-environment systems, we detect *rationality*; as populations, we tend to act in line with utility calculations. Uniquely, humans are partly biological and partly rational. As individuals, we grasp rules (imperfectly), ascribe minds to agents, plan, take part in social institutions and make use of wordings, tools and machines. However, we draw on the resources and skills of populations. How is this to be explained? While bound up with learning to talk (not to mention literacy and numeracy), human agency also uses artifacts and institutions. These perform a dual function both as boundary conditions and as flexible constraints: they serve to measure and also to control. Given the relative predictablity of wordings, we extend our natural intelligence. Accordingly, we now turn to how coordination alters a social actor's sensorimotor and cultural strategies. Coming to act in line with utility calculation depends on learning to concert movements with those of others, exploit available social strategies and, using these, gaining skills in using the artifacts and cultural resources of a community.

Language and Languaging

Since the 18th century human nature has been associated with a mental locus of *ideas*. On such a view, language becomes a transparent conduit between minds

(Locke 1690 [1975]) used to construe verbally-based thoughts (see, Reddy 1979). In the 19th century, this set of metaphors froze as a *code view* that gave us, first, Morse and telegraphy and, later, computers and the Internet. Given the influence of technology, these metaphors were taken up by Saussure and subsequently dominated twentieth century linguistics. However, following Harris (1981), Linell (2005), Love (2004), Kravchenko (2007) and Bickhard (2007) a growing number reject *encodingism*. Far from being a conduit of ideas that are encoded/decoded by minds or brains, language is an ecological whole that functions in many time-scales. It is metabolic or dynamical and, at once, symbolic (Rączaszek-Leonardi 2009). Though part of concerted activity by at least one person, its products are, at once, developmental, historical and evolutionary. Wordings are enacted and, yet, use traditions that are constitutive of the social world. Computers—not living beings—rely on symbolic processes function to encode/decode physical states. Thinking is action: on a machine analogy, total language trains up networks of hypersocial cultural agents that, in their lifetimes, attempt to 'run' languages.

On the conduit view, so-called language 'use' is said to result from the workings of language-systems (e.g., isiZulu, English). Its basis is ascribed to individual knowledge that is represented by a mind or brain. As in Western grammatical tradition, language is described—not around observables (i.e., articulatory activity and pulses of sound)—but by relations between phenomenological forms. Language is thus identified with words, grammars, discourse or constructs that, in some mysterious way, 'reflect' inner thought. Like an artificial system, a brain maps forms onto meaning and, conversely, meanings onto form. Among the problems with any such view is the mereological error of supposing that 'forms' serve brains as input or output. Rather, *people* make and track phonetic gestures that shape how they perceive speech. However, there are no determinate forms *in* the speech wave and we rely on how things are said. As an avalanche of evidence shows, brains exploit rich phonetic information (for review, see Port 2010). Further, we find it hard to track the precise sense of *the words that are actually spoken*. Since connotations affect construal, why does the code view of language persist? Leaving aside the sociology of science, there are two reasons. First, we rely on *the language myth* (Harris 1981): in everyday life and many institutions language is conceptualised in terms of forms that 'contain' messages. Second, inscriptions use writing systems that invite us to think that form is its essence: since language can be reformatted, we view writing as 'like' language—its essence lies in potential reformatting. Many fall prey to what Linell (2005) calls *written language bias*. It is forgotten that, in themselves, inscriptions lack meaning. Like meanings, forms are *abstracta*.

Later, we adduce further reasons for rejecting code views. However, one of the most compelling is that these reify the phenomenological. True to orthodox science, we prefer to begin with measurable phenomena. By starting with speech coordination we commit ourselves to addressing the goal of how this comes to be *perceived* in terms of forms and meanings. Rather than posit a mental or neural locus, this depends on how language spreads across populations. Perception of form and meaning is a multi-agent phenomenon and, thus, language is distributed (see, Cowley 2011a). It is therefore important to distinguish *languaging* from its products (vocal and other

gestures) and the related phenomenology (wordings). Languaging is full-bodied, dialogical, situated and amenable to measurement. Provisionally, it can be defined as "*face-to-face routine activity in which wordings play a part*". Rather than treating forms or meanings as primary, it is recognised that we perceive bodily events as wordings.[2] Emphasis on coordination allows due weight to be given to the fact that languaging predates literacy by tens-of-thousands or years. By hypothesis, all linguistic skills derive from face-to-face activity or languaging. However, it is only over time that children come to make use of these phenomenologically salient and repeatable aspects of second-order cultural constructs (Love 2004). Given the many ways in which they contribute to languaging, they have meaning potential that gives language with a verbal aspect. As Thibault (2011) points out, linguists typically confuse language with second-order constructs. Importantly, in making wordings second-order, we contrast their ontology with that of languaging. Pursuing the contrast, we can use an analogy. Gibson (1979) compared perceiving the world with perceiving pictures. On a card showing Rorschach dots we may see an arrangement of markings and, for example, a dancing bear. Using *discrepant awareness*, we pick up both invariants *of* the picture (e.g., the dots) and invariants *in* the picture (the 'bear'). In languaging too, we pick up invariants of the activity (e.g., how people speak, gesture and use their faces) as well as invariants *in* the activity (e.g., wordings and meanings). Like learning to see pictures, learning to talk draws on discrepant awareness. Just as we see 'things' in pictures, we hear 'things' in utterances. In Cowley's (2011b) terms, we *take a language stance*. On the analogy, this is like learning to see things in pictures while, at the same time, using the body to make one's own (verbal) images. And that, of course, presupposes human agency. Next, we turn to how infants exploit languaging—activity in which wordings play a part—to self-organise their bodies and become human agents who perceive—and make up—wordings.

Human Symbol Grounding

To become fully human, children have to discover how to behave and, among other things, learn how wordings contribute to collective practices. As they become able to play various roles, they benefit from acting and speaking in particular ways. Initially, learning to talk depends on managing concerted action—interactivity—just as much as on wordings. Unlike symbol processors, we use circumstances in coordinating in ways likely to achieve strategic ends. At times we act as others expect and, thus, make what count as valid judgements. Practical skills and shared knowledge shape social action. This, of course, connects ontogenesis, training and education. Next, therefore, we sketch how infants use human symbol grounding to

---

[2]Adults sing, converse, read books, discover new media, and are fooled by advertisements: wordings appear in dreams and silent thoughts. While not languaging, this is also what Love (2004) calls *first-order language*. In all of these activities formal patterns can be said to constrain expressive dynamics.

sensitise to wordings. Whereas babies rely on interactivity, by 3–4 years of age, second-order constructs (and wordings) exert tight constraints on how children act, think and feel. In tracing how the symbols of a community become part of a person, we depend on what Cowley (2007a) terms *human symbol grounding*.

Though the symbols to be grounded consist in more than wordings, these are foundational.[3] Because they are jointly enacted, they link statistical learning, norms and first-person phenomenology. This triple process begins as a baby's brain sensitise to cultural routines. In its first stage, human symbols are grounded into neural networks. Later, infants learn to act in appropriate ways as symbols-for-a-child are grounded into culture. This second stage is further discussed below. Third, once a child's expressive powers develop, she will start to hear wordings: given brains and culture, symbols are grounded into first-person phenomenology.[4] Once wordings shape perception, they serve talk about language or, generally, speaking and acting *deliberately*. Over time, the results drive the functional reorganization of feelings, thoughts and actions. Being able to use wordings deliberately is crucial to rational action. As affect-based co-ordination is supplemented by *the said*, children master new routines (and games). Later, children use special modes of action to structure thoughts (Melser 2004). Much is learned by exploiting context to act epistemically (Cowley and MacDorman 2006). Interactivity enables bodies to use real time adjustments to discover 'organisational' constraints. Though neural predispositions influence ontogenesis, they function through concerted activity. Together, infants and caregivers orchestrate by sensitising to affect marked contingencies. They use co-action or, by definition, how one party used the context of another person's action to come up with something that could not have been done alone.[5] At 14 weeks a mother may be able to use, not touch, but the changing context of her body making her baby fall silent (Cowley et al. 2004). The baby attends to repeated action or, in Maturana's terms (1978), how each orients to the orienting of the other. As a result, circumstances are co-opted in strategic (joint) action. Learning to speak is initially separate from co-action. However, caregivers and infants use the rewards of interactivity to share use of contingencies. As infants manage adult displays, they draw on affect or, in Stuart's (2012) terms, how *enkinaesthesia* prompts us to orient to the felt co-presence of others. Later, these skills become enmeshed with those of vocalising.

In the first stage of human symbol grounding, brains rely on statistical learning. Before birth brains sensitise to rhythms of voices and languages. Infants show

---

[3]Here a symbol can be defined as a cultural constraint that serves in taking the measure or others and/or in controlling one's behaviour: many symbols are prosodic, gestural or enact what Goffman (1959) calls the interaction order.

[4]There is no clear evidence of when this occurs: however, there is abundant evidence that it is based on the skill of making and tracking phonetic gestures (Fowler and Rosenblum 1991; Fowler 2010). Further, since it is necessary to pretending it is likely that children begin to have the necessary experience in the second half of the second year.

[5]Wenger and Sparrow (2007) use experimental work to trace the social and bodily complexity of co-action—and its deep links with our *sense of agency*.

preferential response to the mother's voice (and face) and the *right kind of rhythm* (De Casper and Fifer 1980) and, remarkably, a story heard in the womb (De Casper and Spence 1986). While many animals discriminate, babies have skills in expressive co-ordination. Given rhythmic sensitivity, co-action soon falls under the baby's influence. This was first recognised in Bateson's (1971) work on the protoconversations that reveal 'intersubjectivity' (Trevarthen 1979). Context sensitive co-action is also stressed by Bråten (2007). More recently, the ability to co-ordinate expressive movements (including vocalisation) has been traced to grey matter in the brainstem which, before birth, controls morphogenesis (Trevarthen and Aitken 2001). As motivation develops, contingencies prompt a baby to use the rewards to interactivity to anticipate events. By three months, infants gain skills in controlling vocal, facial and manual expression. Norms already play a part in controlling their enkinaesthetic powers. Language and gesture (not to mention music and dance), thus share a neural basis (Willems and Hagoort 2007). As social actors, we rely on controlling expression in time: an infant uses affect to lock on to the movements by others and, by so doing, engages in dance-like co-action. Even congenitally blind infants move their hands to rhythmic patterns (Tønsberg and Hauge 1996). Those who hear normally, however, use its musical properties to discover the rewards of vocalising.

By three months events begin to show the influence of cultural symbols. Infants sensitise to signs of culture: using coordinated action human symbols are *grounded into culture*. Caregivers use infant gaze, smiles and other expressive gestures as indices of local norms that contribute to co-actional routines. Using both biological tricks and adult displays, infants gain a say in events. While infants and caregivers have *fun* together (see e.g., Stern 1971), affect allows interactivity to build contingencies into dyadic routines or *formats* (Bruner 1983). These help infants with when to initiate, what to expect and, of course, when to inhibit. Surprisingly, a three month old may 'do what its mother wants' by falling silent on command (Cowley et al. 2004); in an isiZulu speaking setting, the baby shows *ukuhlonipha* ('respect'). Dance-like interactivity helps the infant re-enact cultural values. This infant changes parental behaviour in ways that induce learning about situated events. Showing 'respect' (as we describe it), evokes a feeling tone. Without hearing words (or manipulating symbols), the baby comes to value *ukuhlonipha*. Given co-action, cultural contingencies connect with adult display. In this aspect of human symbol grounding, infant motivations exploit adult *experience*. Even if early normative behaviour uses biology (and neural systems that enable adults to shape infant expression) this is a developmental milestone. Before babies learn to reach for objects, caregivers will sometimes act as if their infants 'understand what is said'.

As symbols are grounded into culture, a 3–4 month is increasingly adjudged in terms of how well (or badly) she behaves. Given contingencies and rewards, she sensitises to *circumstances*. Instead of needing stimuli or cues, co-action changes how activity is modulated. For many reasons, the focus of development then switches to learning about objects. Late in the first year, however, the child discovers *secondary intersubjectivity* (Trevarthen and Hubley 1978) during Tomasello's (1999) *9 month revolution*. Bringing social and manipulative skills together, mediated or triadic behaviour emerges. Since language is co-ordinated activity, there is no need for the

identification or recognition of inner intentions. Rather, it is sufficient that adults respond *as if* actions were representational. Infants use contingencies (and compressed information) by acting in ways that *seem* intentional. For example, Cowley (2007b) describes a mother who gets a nine-month old to fetch a block. Far from using inferences, the baby co-ordinates with maternal actions that include shifts of gaze, vocalising 'fetch' three times and using her whole body as a pointer. Fetching thus appears (and, perhaps, feels) intentional thanks to how the mother's vocalisation ('fetch') encourages norm-based activity. Further, if the child can mimic the sound (fetch), this opens up what Tomasello (1999) calls role reversal imitation. This is facilitated by independent concerns that include infant pleasure in self-produced vocalisations. As babbling shapes articulation, by 12 months, a baby 'repeats' syllables. Intrinsic motivations unite with skills and anticipated reactions. In the second year, a baby grasps 'facts' linking the normative with the physical. At times, she may do what is wanted. Once a toddler, she follows commands or, strikingly, directs adult attention and actions. She grasps and utters (what adults hear as) 'words'. As wordings fuse with first-order activity, human agency emerges.

As an infant begins to walk, she is becoming to adopt social roles. While far from reasoning, she draws on virtual patterns and social norms. Given a simple toy, a 12 to 18 month old will enact cultural expectations. In an unpublished study, a French and an Icelandic child-mother dyad played the 'same' game together over several weeks (Sigtryggsdóttir 2007). Whereas the French dyad used this in having fun, the Icelandic partners treated it as goal directed activity. Each baby learned how to elicit rewards. Strikingly, when the Icelandic mother failed to participate, the baby would sometimes self-applaud. Plainly, she exploited—not just affect—but (shared) goals. She has become attuned to the values of her world. Co-ordination enables both parties to use maternal displays of cultural values to organise activity. By participating in routines based in local ways of life, a child learns about fun as well as rationality. Far from relying on sound-patterns, the baby uses rewards that co-vary with what is intended (e.g., 'show respect!'). Quite unknowingly, the child orients to other-directed functions of caregiver verbal patterns. Yet, no one year old *hears* wordings. It is only later that utterances come to be heard as utterances *of* patterns. Early on, first-person experience arises in dynamics of co-action (Cowley 2011a, 2011b). Time passes before a child discovers the potential advantages of using wordings as if they were things.

Coming to *hear* verbal patterns changes human (inter) action and perception. While its neural basis lies in tracking phonetic gestures that feature rich variability, adults use perceptual images as 'words'. All of us have some experience of what is called 'private' thinking by means of public symbols. The skill appears in Piaget's *symbolic stage* when, not by coincidence, infants discover pretending. By hearing more than sound, they discover a 'magical' aspect of language. A two year may say (to a banana), *"hello, who's there?"* Without being able to *hear* telephone talk (a remarkable cognitive skill), such pretending could not arise. Given this perceptual learning, a child learns both to get others to do what *she* wants and to use self-directed speech to shape her action. Whereas language sensitive bonobos can follow novel instructions like 'go and put the orange in the swimming pool' (Savage-Rumbaugh et al. 1998), children excel in different skills. Given their biases, they use

wordings as social resources. Thus, unlike bonobos, humans share attention for its own sake. By age 3 a human child will not only follow instructions but will language differently with peers and in pre-school. She will choose when not to conform: for a child wordings are social affordances. This is just as important in the life of a social actor.

Just as children are not born talking, they are not born rational. Rather, the skills that shape language and reason arise as we identify with aspects of what we hear ourselves say. Co-action prompts infants to orient to local modes of expression. Given a developmental history, layers of agency accumulate together with a history of decision-making. As we redeploy neural resources, we draw on biologically guided interactivity. We learn from a history of anticipating how others will use norms. To act as sophisticated agents training must hone our biological capacities. Human symbol grounding makes us into norm-recognising agents not unlike symbol processors. This depends on individually variable skills in using the language stance to manipulate wordings. Indeed, without this combination, communities would be unable to identify themselves as speakers of a specific language. Without being able to describe words (and rules) as entities that pertain to an autonomous system (e.g. English): we would not believe in abstractions like *minds, selves or societies*.

Using the Resources of Reason

Public language permits *objectively valid* judgements. For Craik (1943), this is exemplified by when, in bridge-building, language is put to symbolic *use*. How this is conceptualised is, of course, a theoretical matter. Since Craik ascribed this to the brain, he viewed language itself as representational. However, the distributed view offers a parsimonious alternative. People learn to refer: they depend on connecting talk about language with languaging (i.e., activity). We learn to take the perspective of the other while linking articulatory patterns with items of shared attention. Once we begin to take a language stance, we hear wordings as wordings that serve to pick out things as things. With practice, we learn to refer or, in other terms, how languaging can be used to pick out objects, persons, events etc.

The skills of a rational human agent depend on both real-time coordination and using the language stance to exploit cultural resources. Given the phenomenological status of wordings, they can be used both literally and in fun. This is because, since they arise from interactivity, they are integrated with action and expression as we enact relationships. In contrast to the fixity of computational symbols, wordings gain effectiveness from flexibility and vagueness. Sense-making arises as they are jointly crafted as persons orient to circumstances and each other. Unlike symbols used in computers, they bear on what people are doing. As part of language flow, interactivity or, in lay terms, *how* we speak is meaningful. Rightly, therefore, many contrast language with man-made codes. Unlike Morse, for example, language is neither monological, disembodied nor dependent on design. Unlike man

made codes, language is dialogical (Linell 2005, 2009). Further, given its embodiment, brains ensure that language is enmeshed with action (e.g., Willems and Hagoort 2007). As one thread in coordination, its literal or denotational meaning is often marginal (Maturana 1978). Circumstances add to how, as living bodies coordinate, we make and construe linguistic signs (Love 2004). Dynamics make language irreducible to words, utterance-types, usage-patterns and so on (Harris 1981; Love 2004). As argued here, people—not language—exploit acts of reference: representationalism is not necessary to cognition (e.g., Anderson 2003; Thompson 2007; Stewart et al. 2010). Indeed, computing systems face the *symbol grounding* problem (Harnad 1990): computations are meaningless to machines. Worse still, where grounded by design, symbols fail to pick out facts. No (currently imaginable) robot could 'know' that, 'Put out the light' is irrelevant to, say, what is in the fridge or a US president's concerns (see, MacDorman 2007). They face *the frame problem* and, for this reason, robots are increasingly designed in ways that enable them to use people to connect with the world.[6]

In contrast to symbol processing, public and collective behaviour enacts skilled co-ordination. Even infants use actions and utterances as representations (e.g., in pointing). During co-action, adults treat infant doings as intentional. Where the infant identifies relevant features, repetition shapes conventional behaviour. For the adult, the infant acts representationally. To extend its cognitive powers, therefore, the baby tracks contingencies. Indeed, as the cultural world increasingly aligns to joint behaviour, the baby learns about co-action. Later, infants come to *hear* words by using interactivity to track contextual indices of local norms. Further, this is a likely basis for using wordings representationally. By using a language stance they become thing-like entities that sustain *belief* in virtual constructs (Cowley 2011a, 2011b). Accordingly, they can be used both as expected and in transgressive ways. Indeed, both approaches lead to sense-making because of how the results are integrated with coordination. Once we perceive second-order constructs as thing-like, we can generate *thoughts* by modulating how we speak or use a pencil to make pictures. Although partly verbal (and symbolic), interactivity connects bodies, activity and cultural experience. Unlike Morse (or computer programs), language is dialogical, multi-modal and realises values (Hodges 2007). Eventually, wording-based reality links with the resources of an interpretative hermeneutic community.

The spread of language prompts social actors to reproduce society. Within a cultural space, children use interactivity to grasp how symbols serve social action. By coming to anticipate how others speak, they discover Wittgenstein's *agreements in judgement* (1958: §242). Using coordination, they develop social strategies that depend on connecting words, circumstances, and the music of expression. In this way, unmeasurable virtual patterns take on cognitive power. Like other living things, we depend on compressed (Shannon) information. In Dennett's (1991) terms, humans

---

[6]One of the most remarkable facts about robots is that, already, they use human consciousness: this is exemplified, for example, when they learn to discriminate what we—but not they—see as colours (see, Cowley 2013). Building on this view, it can be argued that robots are of importance as linguistic informants (Cowley 2008).

use *real-patterns* that include not only cases like gravity and colours but also word-ings. Because of phenomenological status, we can use for example, as ensembles of norms that reflect on other people's expectations. Davidson's (1997) view of the role of thought and language is similar. He proposes a uniquely human *framework* (p. 27): "The primitive triangle, constituted by two (and typically more than two) creatures reacting in concert to features of the world and to each other's reactions, [...] provides the framework in which thought and language can evolve. Neither thought nor language, according to this account, can come first, for each requires the other".

Others concur that thought and language co-emerge from interactions. For Mat-urana (1978), an agent's sense of self fuses with verbal patterns: *structural cou-pling* allows new-borns to engage with caregivers. Their languaging soon becomes oriented to *types* of circumstance (and thus a *consensual domain*). This generates (observer dependent) opportunities for sense-making. Gradually, however, perceiv-ing, feeling and acting are integrated with normative aspects of language. As neu-ral functions change, individuals become *speakers*. In our terms, experience allows discrepant awareness to shape skills based on taking the language stance. A child's sense of self uses coordinated action to link cultural resources with individual skills. For example, we talk about talk, develop narratives, and make up autobiographical memory. By using discrepant awareness, we link circumstances with the past, the possible and the future. It is not the brain but, rather, languaging that underpins refer-ence. Even bridge-building integrates symbolic, practical and skill-based knowledge based on a life history of games that make us (more) rational. Using standardisa-tion, dictionary writing and education, (increasing) weight falls on literal meaning. As this becomes familiar, the language stance favours a detached 'point of view' and more body-centred control of thoughts, feelings and actions—provided that we reproduce social 'reality'.

Human Agency Naturalised

In naturalising human agency, we claim that experience of co-ordinating shapes our cognitive, social, and linguistic skills. Thereby we reformulate Durkheim's old claim that the social explains the social, namely by explaining how biological mem-bers of sapiens develop the dimensions of sociological agency. While bodies are pre-adapted for cultural learning, interactivity prompts brains to compress informa-tion by orienting to verbal patterns, artifacts and norms. We refer by calling up the past, the possible and the future. This is, of course, dependent on institutions and artifacts. Social relations thus underpin reasoning and, of course, skills in making what count as objectively valid or wise judgements. Though we use the results to model social phenomena, their basis lies beyond the brain. We depend on coordi-nating spontaneously while making judicious use of the language stance and the resources of reason. Though languaging retains its importance to face-to-face think-ing, in many other settings, weight falls on treating wordings as wordings. As Piaget

([1962](#)) shows, we come to grasp games of marbles or, later, take part in literacy practices. We increasingly use the language stance to participate in the assemblages that enact joint projects. Human agency is partly eusocial. Its develops from a kind of *fission*: as biological infants become persons, a chain of interactivity transforms what they can do. As this happens, they increasingly discern uses for cultural resources that serve in both individual and collective endeavours.

The transformatory power became especially clear when a bonobo chimpanzee, Kanzi, was raised in a human-like environment (see, Savage-Rumbaugh et al. [1998](#)). Not only did he gain from computer access to verbal resources but these were coupled by close attention from human carers. Bonobo symbol grounding made Kanzi strikingly (partly) human-like.[7] The case contrasts with Davis's ([1949](#)) description of Anna, whose first years of life lacked social embedding and emotional care. She did not speak, could not eat on her own and never laughed or cried. Lack of human company deprived her of opportunities for learning from how people use co-action in orienting to social norms. She never used interactivity in feeling out a cultural world and, as a result, failed to develop the cognitive powers used in social life. Unlike a normal human actor (or Kanzi), her actions were loosely constrained by culture. In short, sociological agency arises as language becomes a dimension of the person. Eighteenth century tradition wrongly plucked words from the world. Language is no transparent medium because, contra Pinker, wordings are not located in the mind (or brain). Rather, they are part of public activity between people, activity that allows even a 14 week old to use co-ordination to show ukuhlonipha ('respect'). The baby does not 'encode' meanings or propositions but, rather, learns from the routines of everyday co-action.

A Sketch of Social Fission

For the social sciences, interactivity and languaging are conceptually important. Although everyday language may be the necessary basis for modelling macro-social phenomena, it seems inappropriate to the micro-social domain. The models of social actor theory (Boudon [1981](#); Coleman [1990](#); Hedström [2005](#)), like those of code linguistics and the computational theory of mind, ignore the world of embodied, conscious beings. In appealing to social fission, we thus naturalize how the social grounds the social. Rather than treat genes and brains as the origins of reason, we argue that children use interactivity to develop locally appropriate kinds of agency. They draw on experience and, crucially, use the language stance to grasp how people, circumstances and situations vary. Brains and genes predispose infants for cultural learning that, by hypothesis, depends on compressing (Shannon) information. They learn about social (and other) affordances as coordination produces experience with norms, artifacts and wordings. Indeed, the flaws in individual rationality speak

---

[7]This depends on the observer's point of view: in many respects, Kanzi remains distinctly a bonobo.

strongly *against* ontological individualism. Rationality derives from social relations: it is a feature of the cultural and institutional environment that drives biological humans to make imperfect use of (what count as) objectively valid judgements.

Since symbolic models capture macro-social patterns, biological humans *discover* the resources of reason. Our agency is made and not born; it emerges from both the physical world and affordances such as languages, artifacts and social institutions. Far from centring on a body (or brain), it depends heavily on how languaging enacts social relations. Though, often, we cannot be literal, judicious use of the language stance brings rich rewards. Combined with appraisal and interactivity, we unearth the value of cultural resources. While sometimes acting individually, joint projects tend to dominate our lives: the artificial matters greatly to human agency. It is thus to be expected that coordination serves to make strategic plans. Following Darwinian logic, it is not at all surprising that social affordances are selected as a result of enacting social relations. This may be why most cultures develop, for example, ways of displaying and recognising kinds of *trust* and *reciprocity*. Interactivity gives rise to a selection history that links up languages, institutions and social norms. Human agents develop intuitive or expert skills alongside those based on the resources of reason. Since human *nature* is so flexible, it is an error to use 'Hobbes's Problem' as evidence for the difficulty of coordination. That said, our limited rationality does create practical problems of aggregation (Spiekermann 2013; Ben-Naim et al. 2013). Rather than view this as a symptom of inherent selfishness, it shows that humans need complex resources that provide results as they move in and out of social aggregates. These make human cognition partly eusocial—much depends on collective modes of action that link the artificial domains of languages, artifacts and institutions.

Interactivity in Human Agency

By acknowledging that cognition cannot be explained by processes within the brain, we move towards a new sociocognitive science. Human agency is constantly re-enacted as interactivity links us with the world. As we do so, we move in and out of social aggregates that draw on languages, artifacts and social institutions. We find our way through the wilds, talk and, for that matter, use computers and develop skills in flying planes. Human agency is not to be identified with the agent. Since it derives from a history of engagement with the world, agency can be traced to four sets of constraints. First, as physico-chemical systems, we exert (and suffer) physico-chemical effects. Second, as living beings, the boundary conditions of our lineage shape the parameters that result from growth, action, learning and development. Third, as human agents, we develop biophysical skills that exploit artificial constraints associated, above all, with artifacts, languages and institutions. Rather than function as boundary conditions, these flexible resources allow us to pursue individual and collective endeavours. Finally, as living subjects, we make and construe artificial affordances. Thus, we are not social actors, languages are not codes

and minds are not symbol processors. In rejecting all such organism-centred views, the distributed perspective holds out the prospect of reintegrating biosemiotics, cognitive psychology, linguistics and the social sciences. The core idea is that our becoming can be traced to interactivity that links agents in larger aggregates within a common world. Although creativity gives rise to artifacts, inscriptions and public performances, its basis lies in how biosocial agents mesh temporal scales while using interactivity. Remarkably, it seems that a single sociocognitive system enables brains, languages and societies to conspire in prompting human bodies to make partial sense of the world. This is crucial to the goals of the field. On the one hand, as noted above, we need to clarify how people come to hear and exploit wordings. On the other, this opens up the much broader question of how phenomenological experience links with organisation in other time-scales. In short, what can be described in language must be traced, on the one hand, to the rapid time-scales of interactivity and neural processes and, on the other, the slow scales that allow groups to differentiate in ways that drive cultural selection. It is there that fission prompts individuals to become the persons that we are. Just as slow scales constrain faster ones, the rapid processes of interactivity and languaging engender human agency—agents who ceaselessly re-evoke the past to explore the adjacent possible.

References

Anderson, M. L. (2003). Embodied cognition. *Artificial Intelligence*, *149*, 91–130.

Bateson, M. C. (1971). The interpersonal context of infant vocalization. *Quarterly Progress Report of the Research Laboratory of Electronics*, *100*, 170–176.

Ben-Naim, J., Bonnefon, J. F., Herzig, A., Leblois, S., & Lorini, E. (2013). Computer-mediated trust in self-interested expert recommendation. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: interactivity, computation and human artifice*. Dordrecht: Springer.

Bickhard, M. (2007). Language as an interaction system. *New Ideas in Psychology*, *25*, 171–187.

Boudon, R. (1981). *The logic of social action: an introduction to sociological analysis*. London: Routledge.

Bråten, S. (2007). Altercentric infants and adults: on the origins and manifestations of participant perception of others' acts and utterances. In S. Bråten (Ed.), *On being moved: from mirror neurons to empathy* (pp. 111–135). Amsterdam: Benjamins.

Bruner, J. (1983). *Child's talk: learning to use language*. New York: WW Norton Press.

Coleman, J. (1990). *Foundations of social theory*. Cambridge: Belknap.

Cowley, S. J. (2007a). How human infants deal with symbol grounding. *Interaction Studies*, *8*, 81–104.

Cowley, S. J. (2007b). The cradle of language: making sense of bodily connections. In D. Moyal-Sharrock (Ed.), *Perspicuous presentations: essays on Wittgenstein's philosophy of psychology* (pp. 278–298). Houndmills: Palgrave Macmillan.

Cowley, S. J. (2007c). Cognitive dynamics and distributed language. *Language Sciences*, *29*, 575–583.

Cowley, S. J. (2008). Robots: the new linguistic informants? *Connection Science*, *20*, 359–369.

Cowley, S. J. (2009). Language flow: opening the subject. *Cognitive Semiotics*, *4*, 64–92.

Cowley, S. J. (2011a). *Distributed language*. Amsterdam: Benjamins.

Cowley, S. J. (2011b). Taking a language stance. *Ecological Psychology*, *23*, 1–25.

Cowley, S. J. (2013). Naturalizing language: human appraisal and technology. *AI & Society*. doi:10.1007/s00146-013-0445-3.

Cowley, S. J., & MacDorman, K. F. (2006). What baboons, babies and Tetris players tell us about interaction: a biosocial view of norm-based social learning. *Connection Science*, *18*, 363–378.

Cowley, S. J., Moodley, S., & Fiori-Cowley, A. (2004). Grounding signs of culture: primary intersubjectivity in social semiosis. *Mind, Culture and Activity*, *11*, 109–132.

Craik, K. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.

Cronin, B. (2004). Bowling alone together: academic writing as distributed cognition. *Journal of the American Society for Information Science and Technology*, *55*, 557–560.

Davis, K. (1949). *Human society*. New York: Macmillan.

Davidson, D. (1997). Seeing through language. *Royal Institute of Philosophy Supplement*, *42*, 15–28.

Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy*, *88*, 27–51.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mother's voices. *Science*, *208*, 1174–1176.

DeCasper, A. J., & Spence, M. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior & Development*, *9*, 133–150.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, *4*, 429–452.

Durkheim, E. (1895 [1998]). *Rules of sociological method*. Basingston: Macmillan.

Fowler, C. A. (2010). Embodied, embedded language use. *Ecological Psychology*, *22*, 286–303.

Fowler, C. A., & Rosenblum, L. D. (1991). The perception of phonetic gestures. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 33–59). Hillsdale: Erlbaum.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday Anchor.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, *42*, 335–346.

Harris, R. (1981). *The language myth*. Oxford: Duckworth.

Hedström, P. (2005). *Dissecting the social*. Cambridge: Cambridge University Press.

Hodges, B. (2007). Good prospects: ecological and social perspectives on conforming, creating and caring in conversation. *Language Sciences*, *29*, 584–604.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge: MIT Press.

Järvilehto, T. (1998). The theory of the organism-environment system: I. Description of the theory. *Integrative Physiological and Behavioral Science*, *33*, 321–334.

Järvilehto, T., Nurkkala, V. M., & Koskela, K. (2009). The role of anticipation in reading. *Pragmatics & Cognition*, *17*, 509–526.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.

Kravchenko, A. (2007). Essential properties of language: why language is not a digital code. *Language Sciences*, *29*, 621–650.

Linell, P. (2005). *The written language bias in linguistics: its nature, origins and transformations*. London: Routledge.

Linell, P. (2009). *Rethinking language mind and world dialogically: interactional and contextual theories of human sense-making*. Charlotte: Information Age.

Locke, J. (1690 [1975]) *An essay concerning human understanding*. Oxford: Oxford University Press. Edited by P. H. Niddich.

Love, N. (2004). Cognition and the language myth. *Language Sciences*, *26*, 525–544.

MacDorman, K. (2007). Life after the symbol system metaphor. *Interaction Studies*, *18*, 143–158.

Maturana, H. R. (1978). Biology of language: the epistemology of reality. In G. A. Miller & E. Lenneberg (Eds.), *Psychology and biology of language and thought: essays in honor of Eric Lenneberg* (pp. 27–63). New York: Academic Press.

Melser, D. (2004). *The act of thinking*. Cambridge: MIT Press.

Pattee, H. H. (1969). How does a molecule become a message? *Developmental Biology. Supplement*, *3*, 1–16.

Pattee, H. H. (1997). The physics of symbols and the evolution of semiotic controls. In *Proceedings from the workshop on control mechanisms for complex systems: issues of measurement and semiotic analysis*, Santa Fe Institute Studies in the Sciences of Complexity. Redwood City: Addison-Wesley.

Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York: Norton.

Port, R. (2010). Rich memory and distributed phonology. *Language Sciences*, *32*, 43–55.

Rączaszek-Leonardi, J. (2009). Symbols as constraints: the structuring role of dynamics and self-organization in natural language. *Pragmatics & Cognition*, *17*, 653–676.

Raguso, R., & Willis, M. (2002). Synergy between visual and olfactory cues in nectar feeding by naïve hawkmoths Manduca Sexta. *Animal Behavior*, *64*, 685–695.

Reddy, M. J. (1979). The conduit metaphor: a case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284–297). Cambridge: Cambridge University Press.

Ross, D. (2007). H. sapiens as ecologically special: what does language contribute? *Language Sciences*, *16*, 710–731.

Savage-Rumbaugh, S. E., Shanker, S. G., & Taylor, T. J. (1998). *Apes, language and the human mind*. Oxford: Oxford University Press.

Sigtryggsdóttir, T. (2007). *Guiding attention leads to changes in child's manifest practical understanding*. Unpublished dissertation, University of Hertfordshire.

Spiekermann, K. (2013). Judgement aggregation and distributed thinking. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: interactivity, computation and human artifice*. Dordrecht: Springer.

Steffensen, S. V. (2011). Beyond mind: an extended ecology of languaging. In S. J. Cowley (Ed.), *Distributed language* (pp. 185–210). Amsterdam: Benjamins.

Stern, D. N. (1971). A micro-analysis of mother-infant interaction: behaviors regulating social contact between a mother and her three-and-a-half-month-old twins. *Journal of American Academy of Child Psychiatry*, *10*, 501–517.

Stewart J., Gapenne, O., & Di Paolo, E. (2010). *Enaction: toward a new paradigm for cognitive science*. Cambridge: MIT Press.

Stuart, S. A. J. (2012). Enkinaesthesia: the essential sensuous background for co-agency. In Z. Radman (Ed.), *The background: knowing without thinking* (pp. 167–186). Houndmills: Palgrave Macmillan.

Thibault, P. J. (2011). First-order languaging dynamics and second-order language: the distributed language view. *Ecological Psychology*, *23*, 210–245.

Thompson, E. (2007). *Mind and life: biology, phenomenology and the sciences of mind*. Cambridge: Harvard University Press.

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

Tønsberg, G. H., & Hauge, T. S. (1996). The musical nature of prelinguistic interaction. The temporal structure and organization in co-created interaction with congenital deaf-blinds. *Nordic Journal of Music Therapy*, *5*, 63–75.

Trevarthen, C. (1979). Communication and co-operation in early infancy: a description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech* (pp. 321–347). Cambridge: Cambridge University Press.

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: research, theory and clinical applications. *Journal of Child Psychology and Psychiatry*, *42*, 3–48.

Trevarthen, C., & Hubley, P. (1978). Secondary intersubjectivity: confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action, gesture, and symbol* (pp. 183–229). New York: Academic Press.

Wenger, D., & Sparrow, B. (2007). The puzzle of coaction. In D. Ross, D. Spurrett, H. Kinkaid, & L. G. Stephens (Eds.), *Distributed cognition and the will* (pp. 17–41). Cambridge: MIT Press.

Willems, R., & Hagoort, P. (2007). Neural evidence and the interplay between language, gesture and action: a review. *Brain and Language*, *101*, 278–289.

Wittgenstein, L. W. (1958). *Philosophical investigations* (2nd ed.). Oxford: Blackwell.