



A Cognitive-Semiotic Approach to Agency: Assessing Ideas from Cognitive Science and Neuroscience

Juan Mendoza-Collazos^{1,2} · Jordan Zlatev¹

Received: 1 December 2021 / Accepted: 2 February 2022 / Published online: 12 March 2022
© The Author(s) 2022

Abstract

Following the levels of intentionality and semiosis distinguished by the Semiotic Hierarchy (a layered model of semiosis/intentionality), and the distinction between original agency (without the use of artefacts) and enhanced agency (the prosthetic incorporation of artefacts), we propose a model of an agency hierarchy, consisting of six layers. Consistent with the phenomenological orientation of cognitive semiotics, a central claim is that agency and subjectivity are complementary aspects of intentionality. Hence, there is no agency without at least the minimal sense/feeling of agency. This perspective rules out all artefacts as genuine agents, as well as simple organisms, since it is highly unlikely that e.g. bacteria have any first-person perspective. Using this model, we review and assess recent proposals on the nature of agency from cognitive science, and neuroscience, and draw conclusions on how to incorporate aspects of them within a synthetic cognitive-semiotic framework.

Keywords Agency hierarchy · Agentive semiotics · Cognitive semiotics · Enhanced agency · Phenomenology · Semiotic hierarchy

Introduction: What Is Agency?

As with many philosophical notions, the concept of *agency* is contentious and ambiguous. Although studies of issues related to agency can be traced back to Aristotle (Schauber, 2003), the use of the term in English does not occur before the fifteenth century (Tønnessen, 2015; Schlosser, 2019). Etymologically, the word derives from the Indo-European root *ag-* “to drive, draw out or forth, move” (Watkins, 2000), used in Greek as *ἄγω* [agw] “lead towards a point, lead on”. In Latin it acquired the sense “to drive at something, to pursue a course of action”, as in the

✉ Juan Mendoza-Collazos
juan.mendoza@semiotik.lu.se

¹ Cognitive Semiotics, Center for Language and Literature, Lund University, Lund, Sweden

² Universidad Nacional de Colombia, Bogotá, Colombia

words *agēns*, *agent-*, *agentia*, and the adjective *āgens* “efficient, effective, powerful” (Lewis & Short, 1879). By 1650, the term is used in the sense of “active operation, a mode of exerting power or producing effect” (Harper, 2021). The currently popular sense “establishment where business is done for another” commenced two centuries later.

In a very broad sense, agency can be attributed to any entity entering into a causal relationship with another, leading to notions like “material agency” (Malafouris, 2013), which risks emptying the concept of meaning, given that every single entity in the universe can be endowed with some causal effectiveness. On the other extreme, agency is commonly understood in philosophy as “the capacity to act intentionally, and the exercise of agency consists in the performance of intentional actions” (Schlosser, 2019). It should be noted that here the term ‘intentional’ is used, again as commonly, in the sense of *deliberate*, implying a degree of reflective consciousness. We are left with a conundrum: is agency just about everywhere in nature, or is it a unique property of human actions, and at that for only for some of them, those that are produced with deliberate, explicit goals?

To help answer this question, we may turn to semiotics, since agency and meaning-making according to many scholars are closely related notions (Barandiaran et al., 2009). But the answers we are likely to receive would depend on the kind of semiotic theory we consult, given well-known differences between the “anthroposemiotics” of who see meaning as a product of human minds and cultures (e.g. Eco, 1981; Morris, 1971), and those who extend meaning-making, and even sign-use, to most, if not all, non-human animals, as typical for biosemiotics (e.g. Von Uexküll, 1982; Sebeok, 1999).¹

One possible way to help resolve this dilemma is to turn to a comprehensive semiotic theory that avoids the “bios vs. anthro” dichotomy, at the same time as it places agency at central stage. In his *Agentive Semiotics*, Niño (2015) defines agency with respect to a number of conditions. From the standpoint of the agent, the three basic conditions are those of *animation* (self-movement), *situatedness* and *attention*.² According to this theory, the agent necessarily has a (conscious or unconscious) goal, an *agenda*. This agenda, on its side, depends both on the agent’s prior actions and on the restrictions from the environment. Niño’s approach aims to integrate different levels of agency, from simpler forms with a basic survival agenda, to the cognitive agency of animals capable of self-awareness, to even higher forms typical of human beings.

While inspired by such ideas, we take several steps back, and address agency from the standpoint of *cognitive semiotics*, which is not a particular theory, but a new science (in the broad sense of the term) dedicated to the study of meaning (also

¹ See Sonesson and Zlatev (2009, eds.) for a Special Issue of the journal *Cognitive Semiotics* on the topic of “Anthroposemiotics vs. Biosemiotics”, showing, among other things, that there are different varieties of either, as well as positions such the present one, which is in several respects intermediary.

² Niño (2015) explains in detail each of these conditions, key theoretical concepts such as *emotion*, *enaction*, *context*, *norms*, *actions* and *events*, and how meaning making and agency are intrinsically related. For a summary, see Mendoza-Collazos (2016).

in the broadest sense), combining methods and concepts from semiotics, cognitive science and linguistics, with the help of phenomenology (Konderak, 2018; Sonesson, 2012, 2015; Zlatev, 2015, 2018). In Section 2, we explain some of the basic methodological principles of cognitive semiotics, and how they compel us to combine experiential accounts of the studied phenomenon, with more detached, “objective” ones. Given that our phenomenon of study is that of agency, we propose a general definition that is meant to be applicable to “biology, phenomenology, and the sciences of mind”, to quote the subtitle of Thompson (2007). Consistent with the level-based approach of Niño (2015), but more closely related to a cognitive-semiotic theory known as the *Semiotic Hierarchy* (Zlatev, 2009, 2018; Zlatev & Konderak, *in press*), and recent proposals to distinguish between basic (bodily) and enhanced (with artefacts and signs) agency (Mendoza-Collazos & Sonesson, 2021), we present an *agency hierarchy*, where higher levels presuppose, but go beyond, lower ones.

Armed with these notions, we turn to a review of several recent theoretical proposals on the nature of agency in cognitive science (Section 3) and neuroscience (Section 4). Given that cognitive semiotics is predicated on the “conceptual-empirical loop” (see Section 2), our aim is to be both critical and constructive in this review, collecting insights for possible future trans-disciplinary studies of agency along the way. We spell out some of these along with our conclusions in Section 5.

Cognitive Semiotics and Agency

Cognitive semiotics arose during the past two decades through collaborations of cognitive scientists, semioticians and linguists who were dissatisfied by the limits of their respective fields. Cognitive science, especially in more recent approaches endorsing *embodied*, *embedded*, *enactive*, *extended* (or “4E cognition”) views on the mind (Newen et al., 2018) is certainly broad, but often lacks deeper analyses of meaning making and sign processes, of the kind that are typical for semiotics. The latter, on its side, has often proposed rather general, and empirically under-supported theories, even when attempting syntheses across the traditional sciences (Eco, 1999). Linguistics continues to be locked in debates concerning what exactly is “language”, from narrow computational views (Hauser et al., 2002), to broad “multimodal” views (Vigliocco et al., 2014). While the latter have been much more compatible with semiotics and cognitive science than the former, there has been a regrettable tendency to extend the concept of “language” to any kind of semiotic system, blurring fundamental differences in the material and semiotic properties that distinguish them (Zlatev, 2019).

While there are different approaches and theories within cognitive semiotics, as within any discipline, most of these have to a greater or lesser degree been influenced by *phenomenology*, the systematic study of lived experience (Sokolowski, 2000). The non-dualist but also anti-reductionist and open-minded take of phenomenology on human experience has already resulted in multiple interdisciplinary approaches within cognitive science, such as *neurophenomenology*, which “seeks articulations by mutual constraints between the field of phenomena revealed by

experience and the correlative field of phenomena established by the cognitive sciences” (Varela, 1996, p.347). A related notion is that of *front-loaded phenomenology*, which “allow[s] the insights developed in phenomenological analyses [...] to inform the way experiments are set up” (Gallagher, 2003, p.91). However, cognitive semiotics aims at more than putting phenomenology at the service of cognitive science and neuroscience, as that would imply that the data from empirical studies are given epistemological priority. This is manifestly not the case when human experience is concerned, since consciousness is not an empirical object on par with others given its transcendental status, and cannot be “naturalized”, without extending our understanding of nature itself (Merleau-Ponty, 1968; Zahavi, 2010).

Hence, cognitive semiotics professes two methodological principles. The first is the *conceptual-empirical loop* (Stampoulidis, 2019; Zlatev, 2015) which postulates that any investigation should begin not with assumed theories and the “predictions” emanating from them (as in the standard “deductive-nomological model”), but with careful reflections on the phenomenon under study, and the concepts used to analyse it. Albeit in brief, this is how we began with the phenomenon of agency in Section 1. Given such a preliminary conceptual investigation, we proceed to the empirical side, and review and conduct observations and experiments that are based on the conceptual distinctions made, and emerging theories. The ambition is then to return to the conceptual side better informed than at the onset, preparing the stage for a new turn of the conceptual-empirical loop. The second principle is that of *pheno-methodological triangulation* (Pielli & Zlatev, 2020), which states that every study should explicitly combine methods from all three perspectives: first-person (intuition-based), second-person (social interaction-based) and third person (detached observation, e.g. experimentation), in this particular order of precedence.³ Fig. 1 represents a combination of these two principles, showing at the same time the close relation to similar ideas with famous pedigree, such as the “hermeneutic circle”.

One particular cognitive-semiotic theory that has emerged through the application of these principles is the *Semiotic Hierarchy* (Zlatev, 2018). This is predicated on the proposal that one of the main concepts of phenomenology, *intentionality*—“the pointing-beyond itself proper of consciousness” (Thompson, 2007, p. 22) or “the directional shape of experience” (Ihde, 2012, p. 24)—is *layered*, forming a hierarchy of meaningfulness. The relation between the layers in the hierarchy is the Husserlian notion of *Fundierung*, especially as elaborated by Merleau-Ponty (1962), with “lower levels prefiguring the higher ones and higher ones consolidating and sublimating the lowers ones, but without breaking away from them” (Zlatev, 2018, p. 6).

³ As an anonymous reviewer points out, some considerations of such methodological triangulation have also been taken from an explicitly biosemiotics perspective by Vehkavaara (2002). However, this proposal gives least primacy for the first-person perspective (called “internal experience” and “introspection”), and even the second-person view seems to be subordinate to the third-person perspective, e.g. “In the “phenomenology of the other one”, both the object-agent and the objects of the phenomenon that the object-agent experiences must be possible objects of our external experience.” (ibid, p. 300). This is converse to the primacy of the three perspectives in the pheno-methodological triangulation of cognitive semiotics.

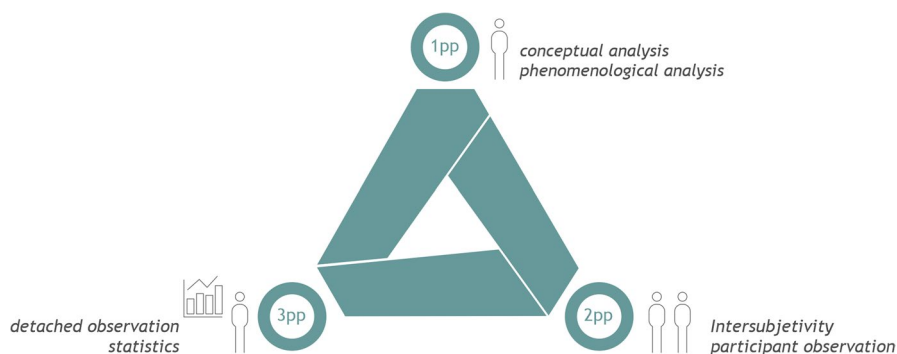


Fig. 1 Combining the cognitive-semiotic principles of the conceptual-empirical loop and pheno-methodological triangulation

To each layer there is a dialectics of free activity (and thus agency) and normativity, with the latter emerging from sedimentations of such activity, thus forming a dialectical spiral of spontaneity and sedimentation (Merleau-Ponty, 1962). The lowest level in the hierarchy is that of *drive and operative intentionality*, and the subject is here the animate and sensing body itself, as argued eloquently by Merleau-Ponty (ibid, p. 296):

There is, therefore, another subject beneath me, for whom a world exists before I am here, and who marks out my place in it. This captive or natural spirit is my body, not that momentary body which is the instrument of my personal choices and which fastens upon this or that world, but the system of anonymous ‘functions’ which draw every particular focus into a general project.

Above this layer is that of (focal) *perceptual intentionality*, which implies explicit (thematic) consciousness and goal-directed attention. These are then preconditions for intersubjectivity, the sharing of experiences between two or more subjects (Zlatev, 2008), and thus for *shared intentionality*, with joint objects of attention and goals. These first three layers of intentionality are more or less accessible to different non-human animals (Preston & De Waal, 2002), but the highest two are uniquely human. According to the theory, sign use is defined as the consciously accessible, denotational relation between expression [E] and intentional object [O], and the *signitive intentionality* corresponding to it is first made possible by bodily mimesis (Zlatev, 2008; Zlatev et al., 2020). On the highest level of the hierarchy is *symbolic intentionality*, made possible by using a highly articulated semiotic system such as spoken language (Sokolowski, 2008), but of course also by signed languages, written languages, and other modern (polysemiotic) media.

As implied by this brief summary, and as argued in the detail by Zlatev (2018), the notion of intentionality is very intimately related to that of meaning-making (semiosis), which should thus also be understood as layered. As Zlatev and Konderak (in press) propose, while intentionality is “outward directed”, from consciousness to the world, meaning-making is a kind of “ricochet” or inverse relation (see Fig. 2), focusing on the experience of this intentionality, from the basic experience

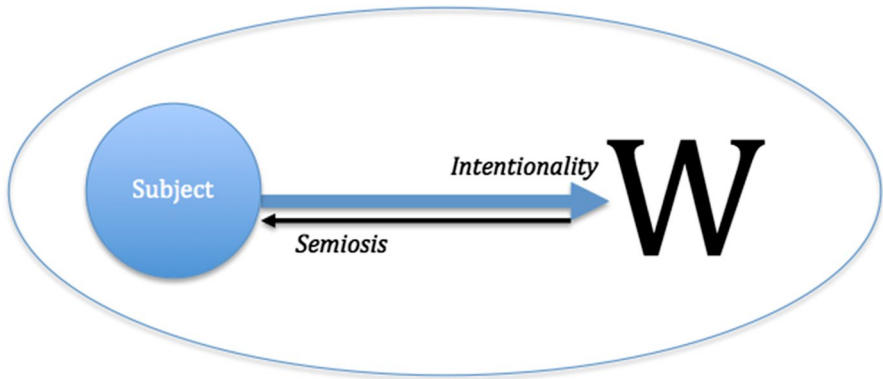


Fig. 2 The reciprocal relation between intentionality and semiosis (meaning-making), adapted from Zlatev and Konderak ([in press](#))

of spontaneous actions, to the interpretations of linguistic texts. Of course, this does not in any way imply that semiosis is “passive”, only that while intentionality always has its apex toward the world, semiosis is what this or that aspect of the world *matters* for the experiencing subject.

How could we understand agency in the context of this approach? Our proposal, remaining to be validated through pheno-methodological triangulation (see above) is that agency is nothing else but *the active, self-generated aspect of intentionality*, which is also why the two concepts (and terms) intentionality and agency are so often conflated. Conversely, *subjectivity* as the irreducible “qualitative feel” of experience (and not the everyday use of the term to mean “bias”), can be seen as the more subject-oriented aspect of intentionality, rather similarly to how Zlatev and Konderak ([in press](#)) proposed understanding semiosis. A similar proposal has recently also been made by a prominent current philosopher of life and mind:

Ordinary ways of thinking furnish us with a number of concepts that help us get a handle on what minds do. One is the concept of *subjectivity*. This concept arrives in a complementary pair with another: *agency*. Subjectivity is a matter of seeming, of for-me-ness. It points to experience as something that happens to a person. Agency is a matter of doing, trying, initiating. Agency is by-me-ness; it is being a source of action and its effects. (Godfrey-Smith, 2020, p. 59)

We find this proposal of a reciprocal relation of subjectivity and agency as mutually complementing aspects of intentionality to be intuitive and consistent with a phenomenological analysis of these notions, and at the same time productive for empirical research. For example, Godfrey-Smith (2020) plausibly proposes that subjectivity/agency emerged first in *metazoa* (animals), with multicellular bodies coordinated by nervous systems, thus proposing a kind of third-person evidence. It is thus consistent with our cognitive-semiotic approach in requiring first-person and third-person evidence for the analysis. The second-person aspect here would consist in whether we make the proposal intuitive for the reader as well. For example, the definition of agency as the active aspect of intentionality, combined with a broad

Table 1 An agency hierarchy: with lower levels standing in a *Fundierung* relation to higher ones, with hypothetical evolutionary primacy, see text for explanations of the examples

Agency type	Level/type of agency	Examples of acts	Examples of agents
Enhanced agency	(6) Mediated by language and other symbolic media	An act of story-telling	A community
	(5) Sign-mediated	Performing a pantomime	A performance troupe
	(4) Artefact-mediated	Making a stone axe	A manufacturer
Original agency	(3) Joint	Dancing in group	A group of dancers
	(2) Reflective	Jumping up to pass over a barrier	A high-jumper
	(1) Operative	Skilled movements without conscious attention, e.g. running	A runner

reading of the latter in accordance with the Semiotic Hierarchy, implies that some “instinctive” reactions, like moving your hand away quickly in response to burning heat, should be understood as minimally agentive: while provoked by an external “stimulus”, the movement away from the flame is experienced as my action, and not as something purely mechanical.

The act of falling after being pushed is not agentive precisely because we experience, at least momentarily, the lack of the active, self-generated aspect of our being-in-the-world (Gallagher & Zahavi, 2012), while all the movements being produced to help ameliorate the fall do exhibit both agency and the sense of it (De Haan & de Bruin, 2010). Such examples are often offered in the literature since they appeal to the basic intuitions of most people. But there is more to both intentionality and agency than bodily actions.

Recently, Mendoza-Collazos and Sonesson (2021) have explicated the notion of *enhanced agency*: the prosthetic incorporation of artefacts to extend the original agentive capacities of “the naked body”: the biological body, from a third-person perspective, or the lived body (*Leib*) from a first person perspective. Enhanced agency emerges from the human ability to design objects that serve particular purposes, thus expanding agents’ ability to act and fulfil both individual and shared goals. Some non-human animals have the ability to elaborate simple objects from natural materials, using processes such as folding, joining, assembling, accumulating, gluing and shaping (Hunt et al., 2006; McGrew, 2013), but the planned improvement of artefacts establishes a fundamental difference between our and other species (Mendoza-Collazos et al., 2021). Further, we need to highlight the difference between making an object and designing an artefact. The ability to design implies innovation and continuous improvement. Enhanced agency is, then, a particular type of agency that is (among other features such as reflective consciousness, signitive intentionality and language, see above) uniquely human, leading some to characterize our species as “the animal that designs” (Colomina & Wigley, 2016).

Merging ideas from the Semiotic Hierarchy and the original vs. enhanced agency distinction, we may propose a layered model of agency such as that shown in Table 1: an *agency hierarchy*, where lower levels are necessary preconditions for

the emergence of higher levels of agency, both conceptually and in evolution, as we indicate below. Given that agency and subjectivity are inextricable from the corresponding levels of intentionality, we require there to be a distinct experience, or a *sense of agency* that corresponds to every layer of the hierarchy. This use of the term “sense of agency” is consistent with – but both broader and more differentiated than – the way it is used both in (phenomenological) cognitive science (Gallagher & Zahavi, 2012), as well as in neuroscience, as we discuss in the following two sections.

Let us briefly discuss examples of the different layers. With respect to the first three layers of *original agency*, displayed bottom-up in Table 1, we can recognize them not only in our own experience, but to some degree also in the behaviour of other animals, using second-person methods (based on social interaction) and third-person methods, like ethology and comparative psychology. The examples we provide below, however, are above all from human experience, given that we appeal to the intuitions of our readers, and hence enact a form of second-person method (see above).

Starting from *operative agency*, the act of running is always purposeful, performed for numerous possible reasons (e.g. hunting, escaping, exercising). Yet, the complex coordination of bodily movements required to carry this out does not require conscious attention, allowing the runner to focus on something else, like dodging obstacles. An explicit decision or “prior intention” to run does not need to have taken place either. As Gallagher and Zahavi (2012, p. 178) point out concerning a similar example:

I might act before I have a chance to decide to act. If, upon approaching the bus stop, I see the bus pulling away, I might start running to catch it. If you stop me and ask, ‘Are you trying to catch the bus?’ my answer would be yes, that was my intention. But it is not clear that I had made any deliberation or conscious decision to run after the bus. I might say, ‘I decided with my feet’, meaning, my decision was in my action, not something separate from it. John Searle (1983) calls this ‘intention-in-action’.

Reflective agency, on the other hand, does require a prior intention, and even the ability to *imagine* the act in question in advance, for example an athlete practicing high jumping. The agent here has to focus their full attention on the target, and (in many cases) to “visualize” the movements they are to perform in order to succeed. In the influential theory of human evolution proposed by (Donald, 1991, 2001, 2013), such imagination is one of the central functions of *bodily mimesis* as a form of consciousness (Zlatev, 2019), before it becomes recruited for sign-based communication, as in pantomime (Zlatev et al., 2020).

Joint agency can be interpreted more narrowly as the product of explicit shared intentions, as when the goal is to carry a heavy object together, or more broadly as the spontaneous intersubjective coordination of movements in a group dance. Both are in their more developed forms human-specific (Tomasello et al., 2005), and may have played a key role in human evolution. But given the complexity of social lives of some other animals like chimpanzees (De Waal, 2007) at least some more simple forms of joint agency in non-human animals cannot be excluded.

On the other hand, we maintain that the three different layers of *enhanced agency* – enhanced step-wise through artefacts, non-verbal signs and language (see Table 1) – are properly conceived of as uniquely human. The simplest of these, *artefact-mediated agency*, is epitomized in the acts of ancient tool-makers when they used natural objects—like stones—as hammers and anvils to produce novel artefacts. While tool-production is the oldest attested evidence of human manufacturing, it is possible that other forms were made with perishable materials. In any case, such manufacturing requires a complex step of procedures that implies complex conscious planning, far exceeding not only intentions-in-action (see the citation above), but also *prior intentions*. Donald (2013) claims again that bodily mimesis as conscious imagination made this possible, and Vaesen (2012) reviews evidence that this capacity remains unique for human beings, despite occasional claims to the contrary.

Even stronger is the evidence for the human specificity, on the population level, of *sign-mediated agency*. This is the case because *signs*, on our understanding of the concept, grounded in Husserlian phenomenology (Sonesson, 2007, 2015; Zlatev, 2018; Zlatev et al., 2020), imply *conscious awareness of the representational status* of certain communicative material expressions (like words, gestures or pictures) or even non-communicative phenomena (like natural symptoms). A key point is that the initial human signs, in both evolution and development (Piaget & Inhelder, 2008) were not yet linguistic, but rather based on similarity (iconicity) or contiguity (indexicality) with their referents. Even non-verbal signs have the capacity to allow the agent greater freedom of reflection and action, by serving as a link between perception and action (Vygotsky, 1978). This has also been observed in semiotically enculturated non-human animals like the bonobo Kanzi and the chimpanzee Pan-banisha, both with the privilege of extensive, and empathetic, training in sign use (Persson, 2008; Savage-Rumbaugh & Lewin, 1994). A clear example of non-verbal sign use that is apparently a human universal is pantomime (Żywczyński et al., 2018).

What language adds to this, and here even enculturated primates falter, is a complex semiotic system of open-class (“categorematic”) signs, denoting objects, events and properties, and closed-class (“syncategorematic”) signs, qualifying the former and helping relate them, which is essential for complex communication and propositional thought (Sokolowski, 2008). Collective narratives such as myths are a natural outgrowth (Collins, 2013; Donald, 1991), and help bond human communities into powerful collective agents (Tomasello et al., 2012). Writing, and eventually media combining language with other sign systems in polysemiotic communication have enhanced human agency even more since then. One example would be the case of modern *corporate identity* and *branding* in which a firm coordinates a full array of communicative media under a unique, pre-established strategy, from the style of written texts to an identifiable style of industrial design (Efer, 2017). We should have this in mind when we consider discussions of the interrelations between agency and design (e.g. Mendoza-Collazos et al., 2021).

To summarize, we have outlined in this section a particular cognitive-semiotic model of agency, formulated a hierarchy that parallels the Semiotic Hierarchy, understood as nested layers of intentionality and semiosis (Zlatev, 2018). A key assumption, with grounds in both phenomenology and biological philosophy, is

that agency is the “other side of the coin” of subjectivity, understood as qualitative phenomenal experience. Thus, there can be no agency without at least a minimal “dosage” of the latter. On other words: *no agency without (at least some) sense of agency*. This is the background against which we review and assess approaches to agency in cognitive science and neuroscience in the following two sections.

Assessing Approaches to Agency in Cognitive Science

We review here three approaches to agency that all reflect recent trends in cognitive science, where the mind is seen as *enactive* in organism-environment interactions, according to which even simple organisms such as bacteria exhibit a basic kind of agency (Barandiaran et al., 2009), or as *extended* beyond the body into artefacts, as in Material Engagement Theory (Malafouris, 2013). As we show below, there are complementary difficulties and advantages to these approaches, from the perspective of the cognitive-semiotic model of agency presented in Section 2. The third approach, integrating phenomenology and cognitive science (Gallagher & Zahavi, 2012) distinguishes several layers of agency, and is both most compatible with the agency hierarchy, and productive for neuroscience studies of agency, as we show in Section 4.

An Enactive Approach to Agency

Varela et al. (1991) were in several respects prophetic some 30 years ago, by predicting that the paradigm of *cognitivism*, based on the concepts of “symbolic” computations and representations (Fodor, 1983), and even *connectionism*, with its “sub-symbolic” ditto (Feldman & Narayanan, 2004), are to be superseded in cognitive science by a paradigm where cognition is redefined in more organic, interactionist and (eventually) experiential terms. The rallying cry for this approach became *enaction*, and the first definition of the latter was: “A history of structural coupling that brings forth a world [...] through a network consisting of multiple levels of interconnected, sensorimotor subnetworks” (Varela et al., 1991, p.206). The origin of this conception lay in the notion of *autopoiesis*, the self-creation of living systems, forming a closed system, and thus a form of identity, while at the same time remaining thermodynamically open and in constant interactions with the environment (Maturana & Varela, 1987). A relevant addition was the notion of *sense-making* (De Jaegher & Di Paolo, 2007) whereby organisms (and perhaps some other systems) actively seek to improve their conditions of self-production, under precarious conditions. Combined with autopoiesis, this “changes the physiochemical world into an environment of significance and valence, creating an *Umwelt*” (Thompson, 2007, p.147), using the term for the meaningful world introduced by von Uexküll (1982). It is on the basis of this conceptual background that Barandiaran et al. (2009) explore the notion of agency in an influential paper. As they point out:

The concept of agency plays a central role in contemporary cognitive science as a conceptual currency across different sub-disciplines [...] amena-

ble to dynamical systems' modeling cutting across brain, body and world and integrating different levels of mechanistic organization into the same explanatory framework. (ibid. p. 1)

But at the same time, they note problems with lack of clear definitions on what exactly constitutes agency and many metaphorical terms, not always recognized as such. Hence, they propose three necessary and jointly sufficient conditions, and consistent with these “a definition of minimal agency” (ibid. p. 11). In general, we find this approach very convincing, and useful for our purposes, even as the authors acknowledge lacunas in their work.

The first and most fundamental condition is that of *individuality* (identity, self-hood), implying a distinction between the agentive system and the environment that the system interacts with from its own point of view, the *Umwelt* mentioned above. This is in contradistinction to objects or artefacts, as argued by the biological phenomenologist Hans Jonas:

[...] to the artifact the identity is accorded; and, insofar as this requires a continuity of memory and tradition in those who do accord it, the identity is the function of another identity, namely, that established in memory, individual and social. This originaive identity of the cognitive subject is a prerequisite for the accorded identity of the object. But this original identity, being that of living systems, is [...] owned by, not loaned to its subject. (Jonas, 1968, p. 239-240, cited by Barandiaran et al., 2009, p. 3)

The second is *interactional asymmetry*, which changes the “structural coupling” provided in the original definition of enaction (see above). On the basis of this:

[...] an agent is a source of activity, not merely a passive sufferer of the effects of external forces. Similarly, an agent is not driven to act by internal, sub-systemic modules, which subordinates the system to the triggering or isolated functioning of a local mechanism. In a sense yet to be properly disclosed, an agent as a whole drives itself, breaking the symmetry of its coupling with the environment so as to modulate it from within. (Barandiaran et al., 2009, p. 4)

Still, there are cases that fulfil these two conditions, but are not truly agentive, for example the tremors of a Parkinson patient. What the person who performs involuntary movements in this manner lacks for this to be a case of agency is the condition of *normativity*, implying that agents actively regulate their interactions with the environment on the basis of “goals or norms according to which they are acting, providing a sort of reference condition” (ibid. p.5). Further, this is intrinsically linked to the *individuality* condition: “component reactions *must* occur in a certain manner in order for the very system to keep going, environmental conditions are *good* or *bad* for the continuation of the system, the system can *fail* to regain stability after a perturbation, etc.” (ibid. p. 7). Thus, this (broad) notion of normativity is more or less equivalent to that of *intrinsic value* (Zlatev, 2003): the agent is not “optimizing” some function attributed from outside, but acting on the basis of their own, possibly internalized, values.

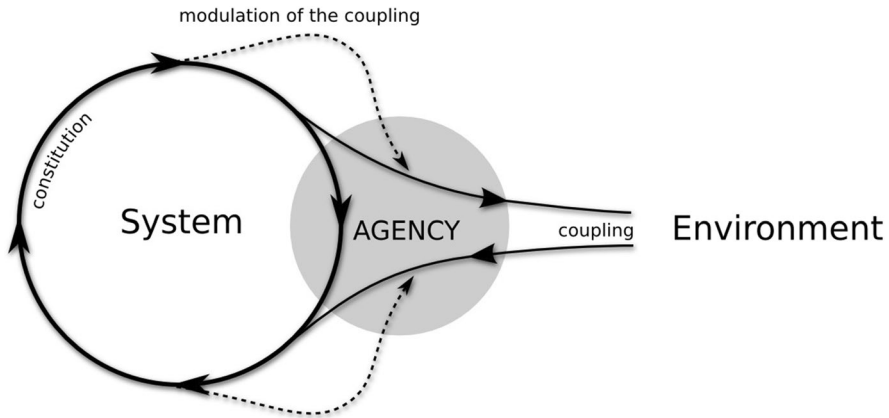


Fig. 3 An illustration of the enactive definition of agency; the system is constituted by a self-sustained network of processes coupled asymmetrically and normatively with the environment. Taken from Barandiaran et al., 2009, Fig. 1

Combined, the conditions are formulated into a “generative definition”, the short version of which holds that: “an agent is an autonomous organization capable of adaptively regulating its coupling with the environment according to the norms established by its own viability conditions (ibid. p. 8) and illustrated in a figure, which we reproduce in Fig. 3.

An advantage of this approach and definition of agency is that it is general, and can apply to bacteria “performing metabolic-dependent chemotaxis” (ibid. p.7), but also to “other domains” (ibid. p. 8), and is not strictly limited to individual living organisms. Thus, it is not susceptible to accusations of “vitalism”. On the other hand, it maintains a clear distinction between agency and causality, and between agents and artefacts, unlike the approach that we discuss in the following sub-section. By understanding agency as a form of “being-in-the-world”, it has at least family-resemblance with our own approach (compare Figs. 2 and 3). It is nevertheless too general from the perspective of our approach, which to remind, requires any degree of agency to be in principle matched with a corresponding degree of subjectivity, a *sense* of agency. Since this is explicitly what Barandiaran et al. (2009) wish to avoid, their “minimal definition” grants agency to bacteria, and possibly even artificial systems, as long as these fulfil the conditions mentioned. But as pointed out by (Thompson, 2007, p. 162), even if supplemented with “sense-making”:

There is no good reason, however, for thinking that autopoietic selfhood [...] involves any kind of intentional access on the part of the organism to its self-making. Second, it seems unlikely that minimal autopoietic selfhood involves phenomenal selfhood or subjectivity, in the sense of pre-reflective self-awareness constitutive of a phenomenal first-person perspective.

Thus, while we do not wish to deny the value of the authors’ “minimal definition” for some domains of study like *artificial life*, we must conclude that it is too minimal to meet our *intuitions* of agency, which require a first-person experience, or

sense, of agency, even in the most minimal manner. Without this, there is no way to perform pheno-methodological triangulation (see Section 2), given that there is no first-person or second-person perspective according to which subjectivity or sense of agency can be attributed to the simplest living organisms, and even less so to their constituents such as organs or cells.

Hence, we would need to supplement this approach with first-person experience, and thus phenomenology, moving from sense-making to *meaning making* (Zlatev, 2018). Barandiaran et al. (2009) are well aware that their approach would require further developments, given that it is focused very much on the “minimal” criteria. For example, it does not deal at all with what we called *enhanced agency*: with tools, signs and language (see Table 1). Toward the end of their paper, they mark this as terrain for future work:

How does niche construction (for example) relate to agency? Should those environmental features that recurrently depend on the agent be considered as part of the agent? What is the status of tools as mediators between agents and environments? (ibid. p.11)

As we show below, some have asked such questions, but reached very different conclusions from those of the enactive approach of Barandiaran et al. (2009).

An Extended Approach to Agency

The move to “extend” the mind beyond not only the brain, but also the body and its “naked” interactions with the world is currently a popular one (Menary, 2010). Despite criticising many tenets of the general extended mind approach, above all on claims of not being radical enough, the Material Engagement Theory developed by Malafouris and colleagues, e.g. (Malafouris, 2013) can be taken as representative. Battling internalism, especially in its cognitivist mind-as-computer variety, and anthropocentrism, the theory endorses the notion of *material agency*, meant to serve as “a wake-up call, for social scientists and archaeologists, to encourage them to consider agency non-anthropocentrically, as a situated process in which material culture is entangled” (Knappett & Malafouris, 2008, p. xii). While Malafouris (2013) adopts a bold form of “methodological fetishism”: “it is more sensible and productive to treat material things as agents (and be wrong) than to deny their agency (and be wrong)” (ibid.: 134), the claim seems to be not that artefacts quite literally have an agency of their own, but rather that agency is a by-product, and emergent property of material engagement.

The theory is also sceptical of other internalist notions like mental imagery and prior intentions, accepting at most intentions-in-action (Searle, 1983)—illustrated in Section 2 with the spontaneous act of running to catch a departing bus—modulated by the material agency of their situatedness. As for sign processes, the approach attempts to define a notion of *enactive signs*, emerging not as representations of what they denote, but through sensori-motor engagements with materials and surfaces,

thus producing, for example, drawings that can only post-hoc be seen as depictions of, for instance, animals on a cave wall.

Comparing this with our cognitive-semiotic approach, but also with the enactive approach of Barandiaran et al. (2009), there is very little in common, apart from the ideological rejection of cognitivism. To begin with, there is nothing corresponding to the “originative identity of the cognitive subject” in sense of Hans Jonas (see above), or the *original agency* of our model. Intentions-in-action are recognized, but apparently only in combination with artefacts. Therefore, it is not clear how Malafouris would treat the agency of the runner, the high-jumper or the dance group (see Section 2, Table 1). In addition, while this is probably not his intention, the theory implies hardly any agency to non-human animals, given that their interaction with objects (and even more making of artefacts) is minimal.

Turning to *enhanced agency*, there are clearly cases where intentions-in-action may predominate, for instance, in trained and sedimented bodily movements required to operate a tool, as the joystick of videogames. But what lends human actions, with or without objects, distinctive creativity is arguably reflective agency, as in sculpting a statue, and joint agency, as in constructing a *maloca*, an Amazonian community house (Mendoza-Collazos & Sonesson, 2021). In fact, there is on-going interaction between prior intentions and intention-in-action in all human, and even some non-human, activities, as can be witnessed in the activities of the chimpanzee Santino who famously collected and hid stones in caches, so that he could throw them at irritating humans when the zoo opened on a later occasion (Osvath, 2009).

In sum, the theory of Malafouris and colleagues, and more generally the “extended mind” approach, might help fill a gap by focusing on enhanced rather than original agency (in our terms), and may serve as an antidote to excessive reliance on “internal” takes on representations and intentionality (in both the philosophical and psychological senses of the term). However, the theory denies, or at least omits, essential properties of agency. One is the fundamentally *asymmetrical* relationship between the subject/agent and the environment, and even the basic *individuality* (identity) of the agent, both (correctly) emphasised as definitional of agency (Barandiaran et al., 2009). Consequently, there is not even the possibility of accommodating our definitional condition: *subjectivity*, or first-person phenomenal experience (Section 2). Thus, the theory denies a central consequence of our approach: that the agency of artefacts is *derived* (Mendoza-Collazos, 2016; Niño, 2015); similar to non-verbal signs and languages, they need to be used, interpreted, or adapted by human beings in human cultures. Thus, the differences between cognitive-semiotic and material engagement approaches to agency far outweigh any similarities (Mendoza-Collazos, 2020).

A Cognitive-Phenomenological Approach to Agency

Following up on the inspirations from phenomenology in the early work on enaction (Varela et al., 1991), but elaborating phenomenological concepts and methods considerably, researchers like Thompson (2007) and Gallagher (2005) have over the past decades helped establish a hybrid field at the intersection of

phenomenology and cognitive science (Gallagher & Schmicking, 2010; Gallagher & Zahavi, 2012). It is this that we refer to as “cognitive phenomenology”, and the topic of agency is often discussed within it, commonly discussed employing the term “sense of agency”. But as Gallagher (2012, *inter alia*) points out, this is a complex notion with many layers.

Initially, Gallagher (2000) distinguished two phenomenological dimensions in relation to movement, both “closely related aspects of minimal self-awareness: *self-ownership* – the sense that it is my body that is moving; and *self-agency* – the sense that I am the initiator or source of the action”. (Gallagher, 2000, p. 16). Often, only the latter is referred to as “sense of agency” (SoA), and opposed to “sense of ownership” (SoO), but barring pathological conditions like schizophrenia, or anarchic hand syndrome, and unless it is someone or something else that has propelled my movement (see Section 2), the kinaesthetic sense of my body in motion implies at least a minimal sense of agency, understood as “a matter of very thin, pre-reflective awareness” (Gallagher & Zahavi, 2012, p. 177). In addition to kinaesthesia, discussed at length by Sheets-Johnstone (2011), there are two aspects to such pre-reflective sense/experience of agency. The first has to do with the urge to move, linked with motor control, and the second with monitoring the effect that one’s movement has produced (Gallagher, 2012). All these pre-reflective experiential aspects of agency correspond to the *operative agency* layer of our model.

But as Gallagher and colleagues (sometimes, but not always) admit, there is more to (the sense of) agency than this, focusing on the goal, and reflecting on the means that would be most appropriate to achieve it. Obviously, this amounts to a layer of *reflective agency*:

For other actions, however, the sense of agency is not reducible to just these embodied and pre-reflective processes. In addition, in many cases we may be reflectively conscious of and concerned about what we are doing. For such actions our sense of agency [...] will be tied to a more reflective sense of intention, by which our attention is directed toward the project or task that we are engaged in, or toward the means and/or end that we aim for. (Gallagher, 2012)

Such reflections do not require language, or any other form of sign use, and hence cannot be a matter of narrative-based *attributions of agency*, of the kind that are assumed to arise as post-hoc rationalisations of our actions by some researchers, inspired by philosophers like Dennett (1991), who claim that self-agency, or even more free will, is a convenient fiction:

[T]he subject’s sense of agency regarding her thoughts likewise depends on her belief that these mental episodes are expressions of her intentional states. That is, whether the subject regards an episode of thinking occurring in her psychological history as something she does, as her mental action, depends on whether she finds its occurrence explicable in terms of her theory or story of her own underlying intentional states (Graham & Stephens, 1994, p. 102, cited by Gallagher, 2012)

As Gallagher and colleagues point out, this could indeed be the case sometimes, for example when participants in experiments use “folk psychological narratives” (Hutto, 2008) to make sense of their own actions, especially when these are manipulated, as in so-called “choice-blindness” studies, more properly referred to as (partial) *manipulation blindness* (Mouratidou, 2019). But these are very special cases, and they can hardly apply to the basic layers of operative agency (in our terms), or even to non-verbal reflections, such as evinced in non-human animals, as the stone caching and throwing chimpanzee Santino (see Section 2).

It is easy to notice the consilience between this approach to agency and our own, given the roots of both in phenomenology. As Gallagher and Zahavi (2012, p.177) stress, “we understand agency, in its proper sense, to depend on the agent’s consciousness of agency”, even if this consciousness is, as pointed out, marginal or recessive. Further, agency (and its experience) is readily admitted to be complex, with dimensions that correspond to layers in our model. Given our summary above, there are at least five different aspects (kinaesthesia, urge, expected effect, reflection and narrative), and three different layers (operative, reflective, language-mediated) of the sense of agency, when conceiving this broadly. And at the same time, we agree that many of these layers are not easily distinguished, since they intertwine and form a composite experience:

Although conceptually we may distinguish between different levels (first-order, higher-order), and neuroscientifically we may be able to identify different brain processes responsible for these different contributories, in our everyday phenomenology we tend to experience agency in a more holistic, qualitative, and ambiguous experience that might be open to a description in terms of degrees of control. (Gallagher, 2012)

As this citation implies, the cognitive-phenomenological approach to agency has implications for the neuroscience of agency, which we address in the following section. But we must comment on a limitation: by focusing on the lowest layers of (the sense of) agency, the theory has little to tell us about agency enhanced with artefacts, signs and language, or even about the joint agency of several people (or other agents) pooling their actions together to form a collective agent of sorts.

Summary

The three different approaches to agency in cognitive science that we discussed have complementary strengths and weaknesses, especially when seen from the perspective of our proposal of the agency hierarchy. The enactive approach of Barandiaran and colleagues contributes with a very useful threshold between minimal agency—on the basis of the conditions of individuality, normativity, interactional asymmetry—and non-agency, the latter characteristic of artefacts. But given that they do not consider the first-person experience/sense of agency as criterial as we do, this threshold is placed somewhat too low.

The phenomenological approach of Gallagher and colleagues, with its distinctions within the experience of agency, is most compatible with the layered approach

of our agency hierarchy. Similarly to our principle of pheno-methodological triangulation, it opens up for operationalizations in terms of third-person methods such as those of neuroscience, as we discuss in the following section. The main lack in this approach was that by focusing so much on the *experience* (sense) of agency within a single person, dimensions (layers) such as joint agency, and agency extended by artefacts, signs and language are largely bypassed—with the possible exception of language-based narratives, for (self) agency attributions.

By rather focusing on artefacts and to some extent on signs (though interpreted in a much too anti-representational way, from our point of view), the de facto extended approach of Malafouris and colleagues (despite their reservations towards the term “extended mind”) can be said to help fill this gap. However, by being so concerned to avoid Cartesian dualism in its quest for interdependent relationship between material things and human beings, this approach misses the strengths of the others: a clear distinction between agency and causality, with artefacts lacking the former, and downplaying the consciousness of agents proper, acknowledging at most intentions-in-action. In sum, we propose that our cognitive-semiotic approach and its agency hierarchy is capable of accommodating the strengths of these models, while bypassing their problems.

Assessing Approaches to Agency in Neuroscience

The Technology, its Problems and Potentials

Cognitive neuroscience has grown exponentially in the past two decades, largely due to the invention of a diversity of advanced techniques for brain imaging like computer tomography, positron emission tomography, ultrasound, and magnetic resonance imaging (Agnihotri et al., 2010; Bradshaw, 1989). Unlike the others, however, *functional magnetic resonance imaging* (fMRI) is commonly applied in studies that are not part of clinical research and the technique most widely used in the neuroscience of agency (David, 2012).⁴ Despite its relatively low temporal resolution, fMRI allows studying both the connectivity and dynamics of brain activity while the person is engaged in particular tasks, which explains its current popularity. This technique uses radiation of strong magnetic fields to manipulate hydrogen molecules of the brain, tracing the activity of neurons, based on signals generated by the blood-oxygen level: more presence of blood-oxygen is assumed to indicate more neural activation (Jenkinson & Chappell, 2018). In this way, experimenters are able to register the activity of the brain triggered by particular stimuli, comparing across conditions or with a baseline condition, usually a resting state with spontaneous neural activity.

⁴ According to David (2012, p. 3): “only a few authors explicitly approached the sense of agency with EEG (Gentsch et al., 2012; Gentsch & Schütz-Bosbach, 2011; Kühn et al., 2011), exclusively focusing on event-related potentials. This classical approach, however, can tell us only little about temporal brain dynamics occurring during agentic processing; rather, it represents a static snapshot of the brain’s activity at the scalp level”.

Although fMRI has many strong points such as high spatial resolution and non-invasive imaging of the tissues in the brain, many researchers within the neuroscience of agency point out numerous methodological and technical challenges (see Braun et al., 2018; Crivelli & Balconi, 2017; David, 2012; Grünbaum & Christensen, 2020; Herdova, 2016; Seghezzi et al., 2019; Zito et al., 2020). For instance, voluntary and involuntary movements of participants (due to discomfort provoked by the scanner) give rise to low-quality data. Other sources of distortions commonly mentioned are materials interference, air-filled sinuses or dental works, which produce inhomogeneities in the magnetic fields leading to miscalculated locations of brain activity. Such errors are said to be “routinely dealt with by the analysis methods” (Jenkinson & Chappell, 2018, p. 21), but statistical corrections are subject of controversy (Brown & Behrmann, 2017). According to Elliott et al. (2020), statistical weakness due to small samples, as well as problems in experimental designs have led to extensive variations, and even apparently contradictory findings across studies.

Some of these problems could presumably be resolved by increasing statistical power and fine-tuning methodologies; these are outside of the scope of our topic, and indeed outside of our competence. However, there are more serious issues with the neuroscience of agency that we believe that our approach can help address such as: (a) *category errors* attempting to literally trace agency in the brain, and (b) *issues of construct validity*, making it unclear what aspect of agency is being operationalized in particular studies, leading in turn to apparently contradictory findings. We illustrate these problems briefly in Section 4.2, while in Section 4.3, we suggest a “reorganization” of the different experimental paradigms used in the field based on the distinctions made in Section 3.3, before summarizing in Section 4.4.

Reductionism and Construct Validity

An implicit and sometimes explicit assumption made by many, albeit not all, neuroscientists is that agency, or at least the “sense of agency” which may very well be an illusion (Dennett, 1991), is to be found in the activity of the brain. For example, in an extensive review of the field, Haggard (2017 p. 206) claims: “The human sense of agency is not a transcendental feature of human nature but the result of specific activity in the brain circuits that underlie voluntary motor control”. A similar view is, for example, expressed by Yomogida et al. (2010, p. 198).

One of the most vocal critics of misguided neuro-reductionism in recent years has been Raymond Tallis, referring to it as “neuromania” (Tallis, 2016). Its main fault is confusing the “the realm of life, meaning and purpose” with that of “cause and effect, mechanisms and physical laws” (Tallis, 2016, p. 155). Given that agency (like subjectivity and intentionality) belongs to the first of these realms, it is what philosophers refer to as a “category error” to look for either identity or (direct) causality between biochemical processes on the one hand, and consciousness and experience on the latter. Tallis is by no means adverse to research elucidating *correlations*, especially when understood as distinct from causation (Aneshensel, 2012), between

neural activity and various states of consciousness. What he opposes are simplistic views of the latter, with multiple references to “illusions of selfhood” ignoring the first-person experiences of human beings. It is, for example, a fundamental fact of such experience that our actions cannot be reduced to a causal chain of physical events, given that each action is not driven by an isolated goal but rooted in a temporal depth, embedded in the sense of past and future. Thus, expressions of agency do not respond to the rules of survival, and cannot be (fully) explained by unconscious activity of the brain controlling the flux of dopamine. As summarized in a similar critique by Fuchs (2017, p. 207):

[...] the mind is not in the brain for it is the overarching manifestation, the gestalt, and the ordered patterns of all relations that we have to our environment as animate beings, and as humans to our fellow humans.

This implies that neural activity should not be understood as the cause of agency, and localizations of greater activity of neurons in certain regions of the brain cannot be equated with (the sense of) agency. Still, while agency cannot be found in neural processes, given that the brain is a highly relevant functional organ of the human body, it is valid to search for correlations between first-person experiential states, and third-person data produced by fMRI studies, as long as we are careful in how we interpret the latter, for example as done within neurophenomenology (Varela, 1996).

Here, however, the second major problem arises: how to define the various theoretical *constructs* of agency in a way that corresponds to the *phenomenon* of agency, i.e. with a high degree of construct validity. In an extensive review, Grünbaum and Christensen (2020) argue that such validity is generally low in the field, since it is not clear what the constructs are, and if a set of experimental procedures is really measuring a given construct or not. Most neuroscientists of agency operate with the two constructs of “sense of agency” (SoA) and “sense of ownership” (SoO), assuming, for example, that stronger *efferent* (“outgoing”) signals generating a movement correspond to stronger SoA, and stronger *afferent* sensory feedback would correspond to SoO. But as pointed out in Section 3.3, it is far from clear that these two constructs correspond to clearly distinct experiences. On the other hand, there are quite different aspects of (the sense of) agency, from the urge to move to attributions of agency that do need to be distinguished before neural correlations are to be sought (see, e.g. Gallagher & Zahavi, 2012, p. 182). What we propose in the following section is to distinguish some of the experimental paradigms used in the field, and to map them to distinctions that have been made in the cognitive-phenomenological approach to agency. This would improve the validity of the various construct of agency in the field, and avoid the problem of contradictory claims on the location of neural correlates of agency.

Improving Construct Validity in the Neuroscience of Agency

Taking into account the five aspects of agency that have been characterized by Gallagher and colleagues (see Section 3.3): kinaesthetic, urge, expected effects,

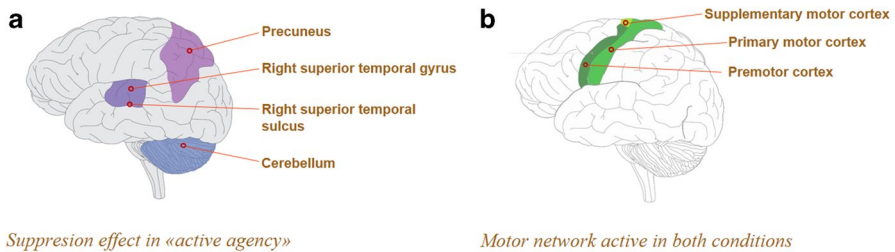


Fig. 4 (a) Brain regions with a suppression of neural activity when participants actively move their own hand. (b) Brain motor network active in both conditions (voluntary movement of the handle and automatically moved). Brain images based on Uhlmann et al. (2020)

reflection, and self-attribution, we can distinguish the following experimental paradigms in the field, each with their corresponding constructs:

Kinaesthetic and Urge Paradigms

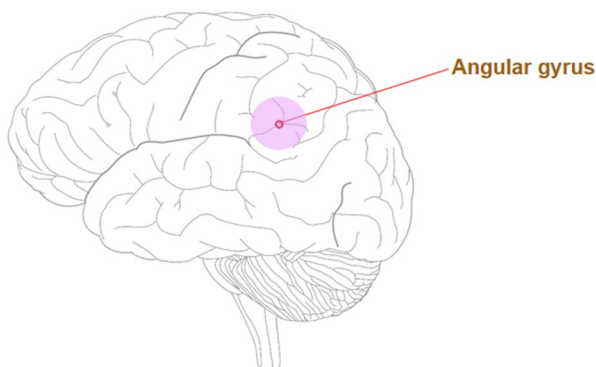
The kinaesthetic aspect of agency is a minimal sense of agency at the pre-reflective level of operative agency. There seems to be little investigations into what the neural correlations of this could be, but it is sometimes said that this is “peripheral experience” relayed by the activity of somatosensory receptors (Haggard, 2017, p. 198).

A clearer sense of agency requires according to Haggard (2017) an internal state of volition or ‘urge’ to move connected to motor control. In a typical paradigm at this level, experimenters ask participants to perform a simple task (e.g. move a handle, press a button, press a key) under two conditions. In one condition, the movement is produced by the participant (so-called “active agency”); in the other condition, the movement is automatically performed by a customized device (so-called “passive agency”). Using this approach, Uhlmann et al. (2020) observed variations in neural processing under the two conditions. Interestingly, the researchers found a reduction of neural activity in the condition of active agency in the areas shown in Fig. 4a, explaining this as follows: “Agency cues consume less resources when participants are able to predict the sensory consequences of their action, that is, when they actively move their own hand” (Uhlmann et al., 2020, p. 2484). Further, they found certain motor-related areas (Fig. 4b) to be active in both conditions, which could possibly be associated with a kinaesthetic aspect of agency: the minimal sense of agency at the most pre-reflective level.

Expected Effect Paradigms

To study the result-expectation aspect of agency, experimenters investigate the neural “processing” associated with action-outcome monitoring: how the brain responds in connection with expected outcomes (Haggard, 2017). Participants are typically asked to estimate the time interval between an action (e.g. pressing a button) and its outcome (e.g. a tone), while researchers compare the neural response to delayed sensory feedback with the response to non-delayed sensory feedback. For instance, van

Fig. 5 Angular gyrus is associated with the detection of delayed feedback during action–outcome monitoring. Brain images based on van Kemenade et al. (2017)



Kemenade et al. (2017) analysed the neural comparator processes during detections of delays. They asked the participants to perform simple button presses, which either produce a tone or present a dot on the screen, or both. The effects of button presses were delayed with a variable amount of milliseconds and participants were asked to judge if they noticed the delay. The findings indicated the importance of a key area, the *angular gyrus*, which functioned as a “supramodal comparator”, given that it was involved in the delayed detection of visual, auditory or combined feedback (van Kemenade et al., 2017, p. 3699, see Fig. 5). Presumably, this can be generalized to stating that the angular gyrus plays a key role in action–outcome monitoring, but more research is needed to be able to conclude this with any certainty.

Reflection Paradigms

Even more cognitive demanding aspects of agency are involved when the agent is focusing on a task requiring conscious attention and reflection. Participants in this kind of experiment are asked to choose among alternatives or to perform an action under different conditions requiring their attention. Neural activity is compared between different choices and conditions. Gertz et al. (2016) applied this paradigm to examine the neural correlates in perception–action couplings and the role of beliefs about the origins of a particular movement. The subjects were asked to track a moving dot on the screen to measure brain activation during oculomotor tracking and were told that the dot could be moved either by a human agent or by a computer program.

By manipulating instructions about the movement (originator and velocity) during this oculomotor tracking task, experimenters searched for the neural correlates of participants' *beliefs about the presence of human agency* on dynamic perception–action coupling: a very different construct compared to those investigated in the studies reviewed so far. The condition where they were told that the dot was moved by a person revealed higher activation in areas such as the frontal pole, the superior frontal gyrus, the anterior cingulate cortex, the posterior cingulate cortex, and in the frontal medial cortex (Gertz et al., 2016, p. 116, see Fig. 6).

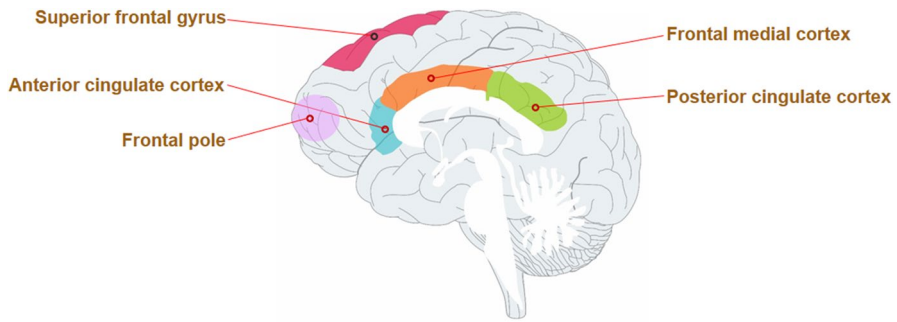


Fig. 6 Areas of the brain with significant neural activations in dynamic perception–action coupling. Brain images based on Gertz et al. (2016)

Attribution-Based Paradigms

Attributions of agency, rather than perceptions or even judgments of agency are more distanced from direct experience, and could in principle be due to an “illusion”, where agency is attributed on the basis of a mistaken belief. In experimental paradigms that seem to target this aspect of agency, participants are sometimes asked to report who is the author of a specific bodily movement, under conditions of ambiguity. For instance, de Bézenac et al. (2016) investigated events in which the cause of an action is not clear or shared between various agents. Participants were asked to perform taps on a keyboard and listen to tones, which could result either from their own hand’s movements or from that of another agent (de Bézenac et al., 2016, p. 67). Participants reported the level of control that they felt, that is, if they thought that the tone was produced by their tapings or by the tapings of others. The finding showed variations of neural processing in subcortical regions such as the *accumbens*, *caudate*, *putamen*, *thalamus*, and the *cerebellum* (see Fig. 7), under the different conditions of agency attribution, e.g. self vs. other.

Summary

Recent neuroscience research on agency seems to be thriving, based on the large number of publications. Yet, as we indicated above, there are still basic technical as well as theoretical problems to be resolved. Admittedly, most researchers are not as naïve as to conflate neural activity (even more, oxygen level variations) with agency and its sense in a one-to-one manner. However, even when investigating “correlations” and “associations”, it is not always clear what theoretical constructs of agency are being used. By distinguishing different aspects of agency with the help of the cognitive-phenomenological approach (Section 3.3) we were able to map these to distinct experimental studies in way that makes it clear that their findings of very different brain areas as neural correlates of agency are not contradictory, since they concern different constructs, and ultimately different phenomena.

Notably, as the level of agentic complexity increases, experimenters seem to include *self-reports* by the participants concerning their experiences of agency,

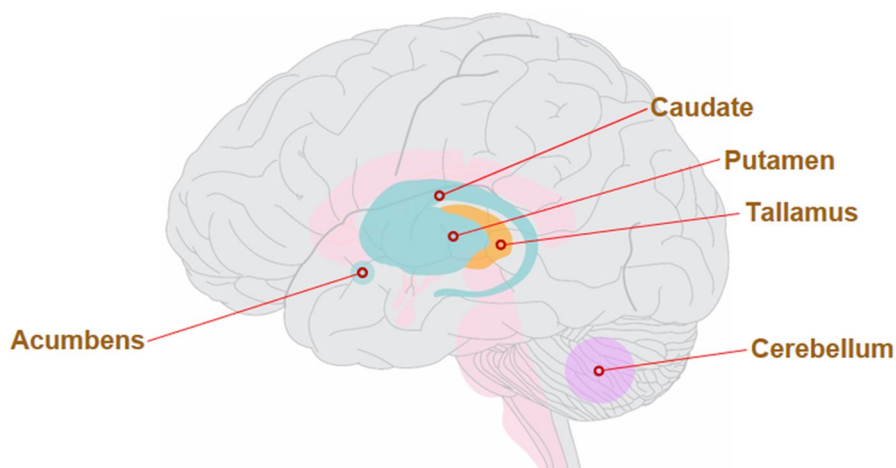


Fig. 7 Neural activations in areas of the subcortical region of the brain and the cerebellum under different conditions of agency attribution. Based on de Bézenac et al. (2016)

which goes at least somewhat in the direction of neurophenomenology. Nevertheless, these “subjective” aspects are generally minimized, and the artificial conditions of the experiments (i.e. over-designed manipulations) are very distinct from everyday life, leading to low ecological validity. Confusing various levels of agency, while over distinguishing the senses of “agency vs. ownership” has made matters even worse. Given that higher levels of agency, related to intersubjectivity and signitive intentions are hardly even explored (Tallis, 2016), we can sum up the current situation in the field by stating an inverse proportion between the complexity of agentive phenomena and the reliability of findings concerning correlations between brain regions with such phenomena (correlational reliability), as shown in Fig. 8.

We would suggest that our cognitive-semiotic approach to agency could help improve this situation. One obvious thing is to distinguish clearly between different aspects/concepts of agency, and to map these to theoretical constructs and their operationalizations more clearly. Another is for experimenters to use pheno-methodological triangulation more explicitly, and investigate both their own, and their participants’ experiences of agency in detail, both before and during the studies. For example, how is the “urge” to move like, or the experiences of either confirmed or disappointed action-outcome relation? How are attributions of self-agency similar and different from those of others?

While emerging technologies such as *hyperscanning* (Misaki et al., 2021; Sadato, 2017) may allow experimenters to study the layer of joint agency, applications of pheno-methodological triangulation would be needed to link these to real life experiences. We propose that neuroscientists can benefit from our approach, and cognitive semiotics more general, to study the complex layers of enhanced agency, asking questions such as: how neural activities differ in cases of intentions-in-action and prior intentions? How engaging with artefacts individually and collectively differ? How thinking with the help of different sign systems correlates with differences in

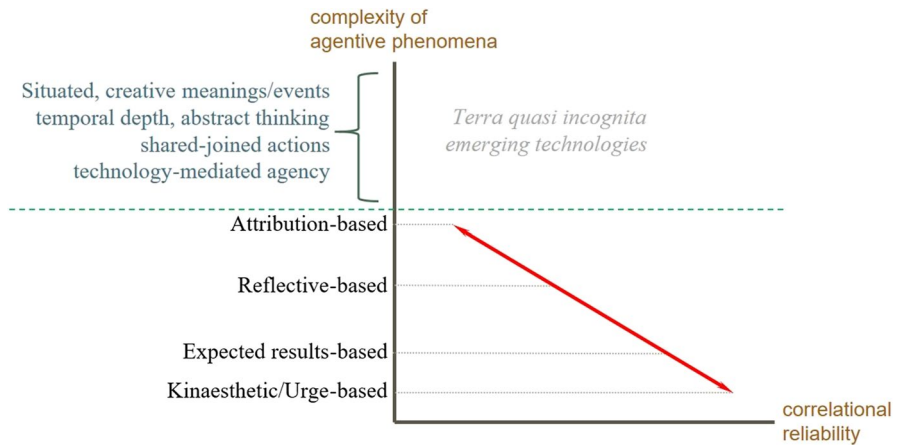


Fig. 8 Reliability of experimental paradigms and the complexity of agentive phenomena: an inverse correlation

agency? Being able to address such questions adequately would more certainly help “straighten up” the red correlational line in Fig. 8, that is, increasing the reliability of studies seeking to establish neural correlates of specific levels of agency.

Summary and Conclusions

We started by pointing out that agency has been understood in either extremely broad or extremely narrow terms, causing confusion with respect to this central philosophical and semiotic concept. With its aim of providing a sort of “middle ground” between ontological and epistemological extremes, the new discipline of *cognitive semiotics* (e.g. Zlatev, 2015) can hopefully help resolve this tension. Compatible with the work of Godfrey-Smith (2020) investigating the biological and evolutionary origins of consciousness, we proposed that agency should be seen as the active, self-generated aspect of intentionality, coupled with subjectivity, the receptive aspect of intentionality, or its “qualitative feel”. This implies that any form of agency must correlate with at least some degree of consciousness: a sense/feeling of agency. This excludes artefacts, inert matter, causality and simpler forms of life as proper agents, reserving agency for animals with unitary bodies and integrated sensory-motor systems. Combining this with the layered understanding of intentionality as in the Semiotic Hierarchy (Zlatev, 2018), and the distinction between original and enhanced agency (Mendoza-Collazos & Sonesson, 2021), we were led to a model of an agency hierarchy consisting of six layers, only some of which imply the capacity for reflective consciousness, and deliberate choice making.

Human beings, while sharing the layers of operative and even reflective agency with other creatures add to these higher levels of agency, and especially those in enhanced agency, which define human uniqueness: technology, sign-use and language-mediated agency. Importantly, such a model does not imply a return to the

hierarchical anthropocentrism of a *scala naturae*, since the relations between layers is dialectical and based on the phenomenological notion of *Fundierung*. Still, defending a particular notion of human uniqueness is necessary despite current fears of “exceptionalism”. As Tallis (2016) warns, denying human uniqueness amounts to diluting human agency not only in its philosophical but also in its political sense: eroding the possibility of channelling our destiny into a better world.

Our cognitive-semiotic proposal of an agency hierarchy follows the principles of cognitive semiotics, given that it was (a) defined on the basis of a conceptual-empirical loop and (b) based on first-person (intuition-based) analysis, second person (empathy, intersubjective corroboration) and, at least potentially, third-person (empirical) methods, including those of cognitive science and neuroscience. This led us to engage more explicitly with approaches to agency in these two fields.

Concerning cognitive science, in Section 3 our model was shown for the most part to be in conflict with *extended* approaches that attribute agency to artefacts, but much more consistent with *enactive* approaches, where agency is clearly distinguished from causality, and based on features such as individuality and intrinsic normativity. Still, given that subjectivity is not included in this list, we found even the enactive approach too inclusive. It is rather the cognitive-phenomenological approach, with its fine attention to different aspects of agency that we found ourselves in most of agreement with.

Turning the approaches to agency from neuroscience in Section 4, we found that some researchers are still subject to naïve views of identity (or causality) between neural and experiential processes, including those of agency. Even those who speak more carefully of correlations are seldom clear about what aspect (or level) of agency is being operationalized, leading to low construct (as well as ecological) validity. Using phenomenological distinctions of agency, we showed examples of at least four different experimental paradigms with target different aspects of agency, making it less mysterious that they find correlations with very different neural processes and brain areas.

It is possible that emerging techniques could potentially avoid some of these problems, as long as they are framed on a refined theoretical approach such as our cognitive-semiotic model of agency. This can expand the scope of neuroscience beyond simple bodily movements in artificial conditions. Our conjecture is that if cognitive (neuro)scientists would adopt a version of our pheno-methodological triangulation more explicitly, they would be able to start with the actual experiences of agency, both their own and those of their participants, and design studies that are more theoretically and ecologically valid. This, on its part, would improve the correlation between reliability and the complexity of agentic phenomena, which is currently far from optimal. Finally, this would allow the cognitive semiotics of agency to engage more fully with third-person methods, and thus to conclude the iteration of the conceptual-empirical loop depicted in Fig. 1.

Acknowledgements We thank the participants of the Cognitive Semiotics Seminar at Lund University for discussions and feedback. We also thank two anonymous reviewers and Göran Sonesson for comments on a previous version. The first author wishes to thank Douglas Niño for valuable feedback and references of emerging technologies. Thanks are also due to Marlon Murillo for helping with the philosophical study of the word “agency” and to Nadia Jiménez for providing the figures of the brain and the graphic design of Fig. 1.

Funding Open access funding provided by Lund University.

Declarations

Conflict of Interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agnihotri, P., Fazel-Rezai, R., & Kaabouch, N. (2010). Comparative analysis of various brain imaging techniques. *Annual International Conference of the IEEE Engineering in Medicine and Biology*, 3029–3032. <https://doi.org/10.1109/IEMBS.2010.5626144>
- Aneshensel, C. (2012). *Theory-based data analysis for the social sciences*. Sage Publications.
- Barandiaran, X., Di Paolo, E., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5), 367–386. <https://doi.org/10.1177/1059712309343819>
- Bradshaw, J. (1989). Brain imaging: An introduction. .
- Braun, N., Debener, S., Spychala, N., Bongartz, E., Sörös, P., Müller, H., & Philippsen, A. (2018). The senses of agency and ownership: A review. *Frontiers in Psychology*, 9, 535. <https://doi.org/10.3389/fpsyg.2018.00535>
- Brown, E., & Behrmann, M. (2017). Controversy in statistical analysis of functional magnetic resonance imaging data. *Proceedings of the National Academy of Sciences of the United States of America*, 114(17), E3368–E3369. <https://doi.org/10.1073/pnas.1705513114>
- Collins, C. (2013). *Paleopoetics: The evolution of the literary imagination*. Columbia University Press.
- Colomina, B. & Wigley, M. (2016). Are we human? Notes on an archaeology of design. Lars Müller.
- Crivelli, D., & Balconi, M. (2017). The agent brain: A review of non-invasive brain stimulation studies on sensing agency. *Frontiers in Behavioral Neuroscience*, 11(229). <https://doi.org/10.3389/fnbeh.2017.00229>
- David, N. (2012). New frontiers in the neuroscience of the sense of agency. *Frontiers in Human Neuroscience*, 6(161), 1–5. <https://doi.org/10.3389/fnhum.2012.00161>
- de Bézenac, C., Sluming, V., Gouws, A., & Corcoran, R. (2016). Neural response to modulating the probability that actions of self or other result in auditory tones: A parametric fMRI study into causal ambiguity. *Biological Psychology*, 119, 64–78. <https://doi.org/10.1016/j.biopsycho.2016.07.00>
- De Haan, S., & de Bruin, L. (2010). Reconstructing the minimal self, or how to make sense of agency and ownership. *Phenomenology and the Cognitive Sciences*, 9(3), 373–396.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507.
- De Waal, F. (2007). *Chimpanzee politics: Power and sex among apes*. John Hopkins University Press.
- Dennett, D. (1991). *Consciousness explained*. Little Brown.
- Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of human culture*. Harvard University Press.
- Donald, M. (2001). A mind so rare: The evolution of human consciousness. Norton.
- Donald, M. (2013). Mimesis theory re-examined, twenty years after the fact. In G. Hatfield & H. Pittman (Eds.), *Evolution of mind, brain and culture* (pp. 169–192). University of Pennsylvania.

- Eco, U. (1981). The theory of signs and the role of the reader. *The Bulletin of the Midwest Modern Language Association*, 14(1), 35–45. <https://www.jstor.org/stable/1314865>. Accessed 17 Mar 2021.
- Eco, U. (1999). *Kant and the platypus*. Secker.
- Efer, O. (2017). Industrial design: The roles and factors of aesthetics, modeling, styling, product brand and branding in design/design education. *Review of Artistic Education*, 13+14, 186–199.
- Elliott, M., Knodt, A., Ireland, D., Morris, M., Poulton, R., Ramrakha, S., Sison, M., Moffitt, T., Caspi, A., & Hariri, A. (2020). What is the test–retest reliability of common task–functional MRI measures? New empirical evidence and a meta-analysis. *Psychological Science*, 31(7), 792–806. <https://doi.org/10.1177/0956797620916786>
- Feldman, J., & Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain and Language*, 89(2), 385–392.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.
- Fuchs, T. (2017). *Ecology of the brain: The phenomenology and biology of the embodied mind*. Oxford University Press.
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14–21. [https://doi.org/10.1016/S1364-6613\(99\)01417-5](https://doi.org/10.1016/S1364-6613(99)01417-5)
- Gallagher, S. (2003). Phenomenology and experimental design: Toward a phenomenologically enlightened experimental science. *Journal of Consciousness Studies*, 10(9/10), 85.
- Gallagher, S. (2005). *How the body shapes the mind*. Oxford University Press.
- Gallagher, S. (2012). Multiple aspects in the sense of agency. *New Ideas in Psychology*, 30(1), 15–31.
- Gallagher, S., & Schmicking, D. (2010). *Handbook of phenomenology and cognitive science*. Springer.
- Gallagher, S., & Zahavi, D. (2012). *The phenomenological mind* (2nd ed.). Routledge.
- Gentsch, A., & Schütz-Bosbach, S. (2011). I did it: Unconscious expectation of sensory consequences modulates the experience of self-agency and its functional signature. *Journal of Cognitive Neuroscience*, 23, 3817–3828.
- Gentsch, A., Kathmann, N., & Schütz-Bosbach, S. (2012). Reliability of sensory predictions determines the experience of self-agency. *Behavioural Brain Research*, 228, 415–422.
- Gertz, H., Hilger, M., Hegele, M., & Fiehler, K. (2016). Violating instructed human agency: An fMRI study on ocular tracking of biological and nonbiological motion stimuli. *NeuroImage*, 138, 109–122. <https://doi.org/10.1016/j.neuroimage.2016.05.043>
- Godfrey-Smith, P. (2020). *Metazoa: Animal minds and the birth of consciousness*. HarperCollins.
- Graham, G., & Stephens, G. L. (1994). Mind and mine. In G. Graham & G. L. Stephens (Eds.), *Philosophical psychopathology* (pp. 91–109). MIT Press.
- Grünbaum, T., & Christensen, M. S. (2020). Measures of agency. *Neuroscience of Consciousness*, 2020(1), 1–13. <https://doi.org/10.1093/nc/niaa019>
- Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 197–208. <https://doi.org/10.1038/nrn.2017.14>
- Harper, D. (2021). Agency. Online Etymology Dictionary. <https://www.etymonline.com/search?q=agency>. Accessed 17 Mar 2021.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *science*, 298(5598), 1569–1579.
- Herdova, M. (2016). Are intentions in tension with timing experiments? *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 173(3), 573–587.
- Hunt, G., Rutledge, B., & Gray, R. (2006). The right tool for the job: What strategies do wild new Caledonian crows use? *Animal Cognition*, 9, 307–316.
- Hutto, D. (2008). *Folk psychological narratives: The socio-cultural basis of understanding reasons*. MIT Press.
- Ihde, D. (2012). *Experimental phenomenology: Multistabilities*. Suny Press.
- Jenkinson, M., & Chappell, M. (Eds.). (2018). *Introduction to neuroimaging analysis*. Oxford University Press.
- Jonas, H. (1968). Biological foundations of individuality. *International Philosophical Quarterly*, 8(2), 231–251.
- Knappett, C., & Malafouris, L. (2008). Material and nonhuman agency: An introduction. In C. Knappett & L. Malafouris (Eds.), *Material agency; towards a non-anthropocentric approach* (pp. ix–xviii). Springer.
- Konderak, P. (2018). Mind, cognition, semiosis: Ways to cognitive semiotics. UMCS.

- Kühn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., & Voss, M. (2011). Whodunnit? Electro-physiological correlates of agency judgements. *PLoS One*, 6, e28657. <https://doi.org/10.1371/journal.pone.0028657>
- Lewis, C., & Short, C. (1879). A Latin dictionary. In G. Crane (Ed.) *Perseus Digital Library*. Retrieved March 31, 2021 from <https://www.perseus.tufts.edu/hopper/>.
- Malafouris, L. (2013). *How things shape the mind*. MIT Press.
- Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human understanding*. Shambhala.
- McGrew W. (2013). Is primate tool use special? Chimpanzee and new Caledonian crow compared. 368 *Phil. Trans. R. Soc. B*. Royal Society of London.
- Menary, R. (2010). *The extended mind*. MIT Press.
- Mendoza-Collazos, J. C. (2016). Design semiotics with an agentive approach: An alternative to current semiotic analysis of artifacts. In J. Zlatev, G. Sonesson, & P. Konderak (Eds.), *Meaning, mind and communication – Explorations in cognitive semiotics* (pp. 83–99). Peter Lang.
- Mendoza-Collazos, J. C. (2020). On the importance of things: a relational approach to agency. *Cognitive Semiotics*, 13(2), 1–11. <https://doi.org/10.1515/cogsem-2020-2034>
- Mendoza-Collazos, J. C., & Sonesson, G. (2021). Revisiting the life of things: A cognitive semiotic study of the agency of artefacts in Amazonia. *Public Journal of Semiotics*, 9(2), 30–52.
- Mendoza-Collazos, J. C., Zlatev, J., & Sonesson, G. (2021). The origins and evolution of design: A stage-based model. In E. Pagni & R. Theisen (Eds.), *Biosemitotics: The natural foundations of symbolism*. Springer.
- Merleau-Ponty, M. (1962). *Phenomenology of perception* (Taylor and Francis e-library, 2005. ed.). Taylor & Francis Group.
- Merleau-Ponty, M. (1968). *The visible and the invisible*. Northwestern University Press.
- Misaki, M., Kerr, K., Ratliff, E., Cosgrove, K., Simmons, W., Morris, A., & Bodurka, J. (2021). Beyond synchrony: The capacity of fMRI hyperscanning for the study of human social interaction. *Social Cognitive and Affective Neuroscience*, 16(1–2), 84–92.
- Morris, C. (1971)[1938]. Writings on the general theory of signs (vol. 16). Mouton.
- Mouratidou, A. (2019). Choice awareness and manipulation blindness: A cognitive semiotic exploration of choice-making. *Public Journal of Semiotics*, 9(1), 1–40.
- Newen, A., De Bruin, L., & Gallagher, S. (2018). *The Oxford handbook of 4E cognition*. Oxford University Press.
- Niño, D. (2015). Elementos de semiótica agentiva. Utadeo.
- Osvath, M. (2009). Spontaneous planning for future stone throwing by a male chimpanzee. *Current Biology*, 19(5), R190–R191.
- Persson, T. (2008). *Pictorial primates: A search for iconic abilities in great apes*. Lund University Press.
- Piaget, J., & Inhelder, B. (2008). *The psychology of the child*. Basic books.
- Pielli, L., & Zlatev, J. (2020). The cyborg body: Potentials and limits of a body with prosthetic limbs. *Cognitive Semiotics*, 13(2), 1–30. <https://doi.org/10.1515/cogsem-2020-2033>
- Preston, S. D., & De Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1), 1–20.
- Sadato, N. (2017). Shared attention and interindividual neural synchronization in the human right inferior frontal cortex. In M. Watanabe (Ed.), *The prefrontal cortex as an executive, emotional, and social brain* (pp. 207–225). Springer.
- Savage-Rumbaugh, E. S., & Lewin, R. (1994). *Kanzi: The ape at the brink of the human mind*. Wiley.
- Schauber, N. (2003). Aristotle on animals, agency, and voluntariness. *Richmond Journal of Law and the Public Interest*, 7(1), 26–50. <https://heinonline.org/HOL/P?h=hein.journals/richlapin7&i=28>. Accessed 17 Mar 2021.
- Schlosser, M. (2019) Agency. In: E. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/win2019/entries/agency/>. Retrieved 1 April 2021.
- Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.
- Sebeok, T. (1999). The sign science and the life science. *Applied Semiotics*, 3(6/7), 85–96. <https://doi.org/10.5840/cpsem19901>
- Seghezzi, S., Zirone, E., Paulesu, E., & Zapparoli, L. (2019). The brain in (willed) action: A meta-analytical comparison of imaging studies on motor intentionality and sense of agency. *Frontiers in Psychology*, 10, 804. <https://doi.org/10.3389/fpsyg.2019.00804>
- Sheets-Johnstone, M. (2011). *The primacy of movement*. Expanded 2nd edition. John Benjamins.
- Sokolowski, R. (2000). *Introduction to phenomenology*. Cambridge University Press.

- Sokolowski, R. (2008). *Phenomenology of the human person*. Cambridge University Press.
- Sonesson, G. (2007). From the meaning of embodiment to the embodiment of meaning: A study in phenomenological semiotics. In T. Ziemke, J. Zlatev, & R. Frank (Eds.), *Body, language and mind. Vol 1: Embodiment* (pp. 85–128). Mouton de Gruyter.
- Sonesson, G. (2012). Between homeworld and alienworld: A primer of cultural semiotics. In W. B. Ernest (Ed.), *Sign culture – Zeichen Kultur* (pp. 315–328). Königshausen & Neumann.
- Sonesson, G. (2015). Phenomenology meets semiotics: Two not so very strange bedfellows at the end of their Cinderella sleep. *Metodo: International Studies in Phenomenology and Philosophy*, 3(1), 41–62.
- Sonesson, G., & Zlatev, J. (Eds.). (2009). Anthroposemiotics vs. Biosemiotics. *Cognitive Semiotics*, 4.
- Stampoulidis, G. (2019). Stories of resistance in Greek street art: A cognitive-semiotic approach. *The Public Journal of Semiotics*, 8(2), 29–48.
- Tallis, R. (2016). Aping mankind. Neuromania, darwinitis and the misrepresentation of humanity. *Acumen*.
- Thompson, E. (2007). *Mind in life: Biology, phenomenology and the sciences of mind*. Harvard University Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(05), 675–691.
- Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation. *Current Anthropology*, 53(6), 673–692.
- Tønnessen, M. (2015). The biosemiotic glossary project: Agent, agency. *Biosemiotics*, 8(1), 125–143. <https://doi.org/10.1007/s12304-015-9229-0>
- Uhlmann, L., Pazen, M., van Kemenade, B., Steinsträter, O., Harris, L., Kircher, T., & Straube, B. (2020). Seeing your own or someone else's hand moving in accordance with your action: The neural interaction of agency and hand identity. *Human Brain Mapping*, 41(9), 2474–2489. <https://doi.org/10.1002/hbm.24958>
- Vaesens, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences*, 35, 203–262.
- van Kemenade, B., Arikani, B., Kircher, T., & Straube, B. (2017). The angular gyrus is a supramodal comparator area in action–outcome monitoring. *Brain Structure and Function*, 222(8), 3691–3703. <https://doi.org/10.1007/s00429-017-1428-9>
- Varela, F. (1996). Neuropsychology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, 3(4), 330–349.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.
- Vehkavaara, T. (2002). Why and how to naturalize semiotic concepts for biosemiotics. *Sign Systems Studies*, 30(1), 293–313.
- Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B*, 369(1651), 20130292.
- Von Uexküll, J. (1982)[1934]. The theory of meaning. *Semiotica*, 42(1), 25–82. <https://doi.org/10.1515/semi.1982.42.1.25>
- Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.
- Watkins, C. (Ed.). (2000). *The American heritage dictionary of Indo-European roots*. Houghton Mifflin Harcourt.
- Yomogida, Y., Sugiura, M., Sassa, Y., Wakusawa, K., Sekiguchi, A., Fukushima, A., Takeuchi, H., Horie, K., Sato, S., & Kawashima, R. (2010). The neural basis of agency: An fMRI study. *Neuroimage*, 50(1), 198–207. <https://doi.org/10.1016/j.neuroimage.2009.12.054>
- Zahavi, D. (2010). Naturalized phenomenology. In S. Gallagher & D. Schmicking (Eds.), *Handbook of phenomenology and cognitive science* (pp. 2–19). Springer.
- Zito, G., Wiest, R., & Aybek, S. (2020). Neural correlates of sense of agency in motor control: A neuroimaging meta-analysis. *PLoS One*, 15(6), 1–17. <https://doi.org/10.1371/journal.pone.0234321>
- Zlatev, J. (2003). Meaning = life (+ culture): An outline of a unified biocultural theory of meaning. *Evolution of Communication*, 4(2), 253–296.
- Zlatev, J. (2008). From proto-mimesis to language: Evidence from primatology and social neuroscience. *Journal of Physiology – Paris*, 102, 137–152.
- Zlatev, J. (2009). The semiotic hierarchy: Life, consciousness, signs and language. *Cognitive Semiotics*, 4, 170–201.
- Zlatev, J. (2015). Cognitive semiotics. In P. Trifonas (Ed.), *International handbook of semiotics* (pp. 1043–1067). Springer.
- Zlatev, J. (2018). Meaning making from life to language: The semiotic hierarchy and phenomenology. *Cognitive Semiotics*, 11(1), 50. <https://doi.org/10.1515/cogsem-2018-0001>

- Zlatev, J. (2019). Mimesis theory, learning and polysemiotic communication. In M. Peters (Ed.), *Encyclopedia of educational philosophy and theory*. Springer.
- Zlatev, J., & Konderak, P. (in press). Consciousness & Semiotics. In J. Pelkey (Ed.), *Handbook of semiotics*. Bloomsbury Press.
- Zlatev, J., Żywiczyński, P., & Wacewicz, S. (2020). Pantomime as the original human specific communicative system. *Journal of Language Evolution*, 5(2), 156–174. <https://doi.org/10.1093/jole/lzaa006>
- Żywiczyński, P., Wacewicz, S., & Sibierska, M. (2018). Defining pantomime for language evolution research. *Topoi*, 37(2), 307–318.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.