JASPER C. VAN DEN HERIK
Erasmus University, Rotterdam, Netherlands

ATTENTIONAL ACTIONS – AN ECOLOGICAL-ENACTIVE ACCOUNT OF UTTERANCES OF CONCRETE WORDS

This paper proposes an ecological-enactive account of utterances of concrete words – words used to indicate observable situations, events, objects, or characteristics. Building on the education of attention model of learning, utterances of concrete words are defined as *attentional actions*: a repeatable form of behaviour performed by a person to *indicate* (i.e. point out) a particular aspect of the current situation *to* someone *in order* to achieve something. Based on recent empirical evidence on categorical colour perception, attentional actions are proposed to constrain the ongoing phenotypic reorganisation of persons into *task-specific devices*. The paper ends by situating the proposed account in a wider theoretical perspective on language. This paper serves two purposes: *first*, it undermines the scope objection against the ecological-enactive approach, and *second*, it provides a novel explanation for recent empirical evidence with respect to the role of language in categorical colour perception.

Key words: attentional action, enactivism, ecological psychology, word learning, word-meaning, categorical perception

The cognitive sciences are currently making a *pragmatic turn* (Engel et al., 2013). The dominant *representational theory of cognition*, where cognition is conceived of as computation over mental representations, is gradually giving way to the ecological-enactive approach, which foregrounds the action-oriented, embodied, and situated nature of cognition (see §1.3 for a short introduction). This pragmatic turn has implications for understanding linguistic behaviour. In this paper, I reconceive the cognitive function of utterances of concrete words, words that concern observable situations, events, objects, or characteristics.

Address for correspondence: Jasper C. van den Herik, The Erasmus School of Philosophy, Erasmus University, Postbus 1738, 3000 DR, Rotterdam, Netherlands. E-mail: jasperherik@gmail.com

There are two reasons for starting with concrete words. First, concrete words are a paradigmatic example in the dominant code-view in the representational theory of cognition, according to which words function as codes for mental representations (see §1.1). Second, the empirical research on the relation between language and cognition on which this paper draws often relies on these concrete words, in particular colour words (see §2).

In this paper, I propose a way to understand utterances of concrete words from the ecological-enactive approach by starting from the *constraint view*. On this view, language is reconceptualised as a mode of action that functions as enabling constraints on cognitive and interactive dynamics. Building on the ecological-enactive conception of attention, I define an attentional action as a repeatable form of behaviour performed by a person to indicate a particular aspect of the current situation to someone in order to achieve something. I argue that a concrete word, whether performed as a single-word utterance or in the context of a grammatically structured utterance, has to be understood in terms of its constraining effect on ecological-enactive attention. In other words, utterances of concrete words are attentional actions.

This paper serves two purposes: first, it undermines the familiar scope objection against the ecological-enactive approach, according to which it can only explain basic cognition and cannot explain 'higher' cognition such as linguistic cognition. Second, it provides a novel explanation for recent empirical evidence with respect to the role of language in categorical colour perception.

This paper is structured as follows: first, I introduce the code view and the ecological-enactive approach. I propose an account of a child's first utterances of concrete words as attentional actions based on the constraint view of language and a discussion of learning language in line with the education of attention model of learning (§1). Then I discuss some recent empirical findings that show verbal colour categories to have an effect on putatively non-linguistic colour perception (§2). In line with this evidence, I propose an account of the cognitive effects of attentional actions in terms of constraints on the phenotypic reorganisation of task-specific devices (§3). In the final section, I situate the account of attentional actions in a larger theoretical perspective on language by means of remarks on the context-dependency and reflexivity of language, grammatical structure, and written language (§4).

1. Words: from codes to constraints

1.1. The code view

Entrenched in our thinking is a picture of language as a carrier of thought. We find articulations of this picture in everyday English expressions such as *putting* a concept *into* words, or a text *containing* new ideas.[1] In theoretical reflections on language, this picture manifests in the idea that words encode

---

[1] Reddy (1993) lists a great number of examples from everyday English.

mental meanings. This has been dubbed the *code view* (Harris, 1990; Linell, 2005; Love, 2004; Kravchenko, 2007). On this view, communication consists in the encoding of mental meanings in public symbols and the subsequent decoding of these public symbols back into mental meanings.

The code view lies at the root of many theories of language. Notable historical proponents of the code view include Aristotle (1975, p. 43), who claimed 'spoken sounds are symbols of affections in the soul', and John Locke (1690/2004, p. 363), who took words to be 'sensible marks of ideas'. De Saussure (1922/2011, p. 66) defined the linguistic sign as a connection between a concept and a representation of a linguistic sound. He described linguistic communication in terms of the speaking circuit: a concept 'unlocks' a sound representation, which is transmitted to the organs used in producing speech, where they are transfused into actual sound patterns which then travel to another person in order to undergo the reverse operation (Ibid., pp. 11–12). The second half of the 20th century witnessed the rise of the *representational theory of cognition*, which enabled the formulation of new versions of the code view. According to one version of this theory, thought itself is language-like, and understanding public language is essentially a code-like translation to a language of thought, called *mentalese* (Devitt, 2006; Fodor, 1975, 2007; Pinker, 1994, 2007). Other versions of the representational theory of cognition rely instead on the broader notion of semantic or conceptual representations. For example, Cairns and Cairns (1976, pp. 17–18) claim that 'The listener, B, must decode A's message by converting the sounds into a semantic representation'. Nowadays, the idea that words encode semantic representations is a staple in linguistics textbooks. In *The Blackwell Handbook of Linguistics*, for example, Cruse (2017, p. 253) states:

> Each of us has in our cognitive system some kind of inventory of all the words that we know, together with all the information – semantic, grammatical, and phonetic / graphic – necessary for their correct use. […] The inventory is accessed via written or spoken forms every time we hear or read something in a language we know, and via some kind of semantic representation every time we produce language.

In *The Lexicon: An Introduction*, part of the *Oxford Textbooks in Linguistics* series, Ježek (2016, p. 5) claims that the 'existence of a word' is the result of 'the direct association of a concept with a lexical form'.

Proponents of the code view thus differ in what they take words to be encoding: concepts, ideas, semantic representations, etc. For the purposes of this paper I shall use *mental representation* as an umbrella term to denote this panoply of mental meanings. The code view can then be defined as follows: A word is a code for (a set of) mental representations. Note that due to synonymy and polysemy, the mapping between words and (sets of) mental

representations cannot be a one-to-one mapping, but instead is a many-to-many mapping (Cruse, 2017, p. 252).

I introduce the code view in order to use it as a foil. The reason for this use is that the code view has had a deep influence on Western theoretical reflections on language and remains the dominant view, although there are of course alternative views.[2] The purpose of this paper is neither to determine how wide-spread the code view is, nor to provide arguments to criticise it, but instead to use it as a contrast. This enables bringing out the defining features of the constraint view to which this paper aims to contribute.

1.2. Mental representations: can't have, don't need

Proponents of the pragmatic turn in the cognitive sciences claim that cognition is embodied. Yet, there is substantial disagreement about what this claim means (Anderson, 2003; Kiverstein & Clark, 2009; Wilson, 2002; Wilson & Golonka, 2013). In its weakest interpretation, an 'embodied concept is a neural structure that is part of, or makes use of the sensorimotor system of our brains' (Lakoff & Johnson, 1999, p. 104). As opposed to *classical cognitivism*, mental representations are no longer amodal. But cognition is still a computational process over discrete mental representations on this weak interpretation. The strongest interpretation, known as *radical embodied cognition*, aims to provide a thoroughly non-representational approach.[3] Note that a non-representational approach forswears *mental* representations. This does not entail forswearing *public* representations such as maps, assertions, diagrams, charts, or flatpack furniture assembly instructions.

The account in this paper is based primarily on two research programmes that espouse a form of radical embodied cognition: ecological psychology (Chemero, 2009; Gibson, 1979; Rietveld & Kiverstein, 2014; Turvey, Shaw, Reed, & Mace, 1981) and enactivism (Hurley, 1998; Hutto & Myin, 2013; O'Regan & Noë, 2001; Stewart, 2010; Noë, 2012; Varela, Thompson, & Rosch, 1991). I therefore speak of the *ecological-enactive approach*. Note that both enactivism and ecological psychology are research programmes within which different theories are espoused, of which I will not provide an exhaustive taxonomy here.

---

[2] Examples of alternatives include Ryle, who does not believe that there are mental meanings (Tanney, 2015), the later Wittgenstein (1953/2009), Voloshinov's (1930/1973) Marxist philosophy of language, Maturana's (1978) biological approach to language, Bakhtin's (1982) dialogism, integrational linguistics (Davis & Taylor, 1990; Davis & Taylor, 2003) and the distributed language perspective (Cowley, 2011a).

[3] Named as such by Clark (1997, p. 148), who defines it as follows: 'Structured, symbolic, representational, and computational views of cognition are mistaken. Embodied cognition is best studied by means of noncomputational and nonrepresentational ideas and explanatory schemes.' Proponents that fly under the banner of radical embodied cognition include Chemero (2009) and Hutto and Myin (2013; 2017). The word radical here should not be understood as a radicalisation of embodied cognitive science, but instead in the sense of having different roots. These theoretical roots include American pragmatism (Peirce, James, Dewey, and Rorty), behaviourism (Ryle and Skinner), ecological psychology (James and Eleanor Gibson), theoretical biology (Maturana and Varela), dynamic systems theory (Thelen & Smith, 1994; Van Gelder, 1998) biophilosophy (Jonas), and existential phenomenology (Heidegger and Merleau-Ponty).

A proponent of the ecological-enactive approach to cognition can employ two lines of argument with respect to the representational theory of cognition: *can't-have* and *don't-need* arguments (Hutto & Myin, 2013). Can't-have arguments are negative. They aim to show that there are currently no naturalistically credible theoretical resources available for accounting for mental representations. Hutto and Myin (2013), for instance, argue that there is no satisfactory account of how mental representations get their representational contents. If a can't have argument is successful, representational approaches to cognition are off the table. However, as Chemero (2013, p. 3) points out, can't have arguments are likely to be met with a 'collective shrug'. That is, these arguments against the representational theory of cognition tend not to convince proponents of that theory. Don't-need arguments, on the other hand, are positive. They aim to show that mental representations need not be invoked in order to explain a particular cognitive phenomenon. If a don't need argument is successful, this does not imply that the non-representational explanation should automatically be preferred over an alternative representational explanation. However, meta-theoretical considerations – such as ontological parsimony and explanatory power – can then be used to mount arguments in favour of either explanation.

Note that the two strategies can be employed independently. It is possible to argue that we can't have mental representations without supplying a non-representational alternative explanation, and it is possible to propose a non-representational explanation without arguing that that we can't have mental representations. The intended contribution of this paper is only to provide a don't need argument. This is done by showing a possible account of utterances of concrete words that does not invoke mental representations, and show how it can account for recent empirical evidence.

1.3. The ecological-enactive approach in a nutshell

The guiding idea of the enactive approach is that perception and action are intimately intertwined. In the dynamic coupling between organism and environment, a feedback loop is set up (see Figure 1). This is called a *sensorimotor loop* (Stewart, 2010). Cognition is defined as *perceptually guided action* (Varela et al., 1991, p. 173), and thus becomes something the organism does in interaction with its environment (Myin, 2016; Noë, 2012). Cognitive processes do not take place inside the head of the organism, but are relational in nature.

Ecological psychologists similarly stress the intertwining of perception and action in the concept of *affordance* (Gibson, 1977). Organisms perceive what the environment *affords* doing, that is, possibilities for action. For example, a flight of stairs is perceived as climbable and a cup is perceived

as graspable.[4] The perceived affordances depend *inter alia* on aspects of the environment, the skills of the organism (not everything that is liftable for a world champion weightlifter is liftable for me), as well as the material possibility we have in exercising these skills, for instance the availability of artefacts. As Stewart (2010, p. 29) remarks, 'a snow-covered mountain becomes an entirely different place if you have skis on your feet (and if you know how to ski!)'.
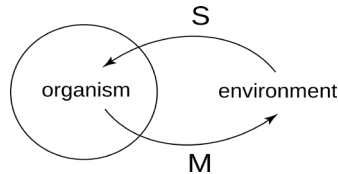

Figure 1. The sensorimotor loop: a feedback loop consisting of sensory (S) and motor (M) processes (based on Stewart (2010, Figure 1.1)).

We never perceive only one affordance. An individual always perceives a *field of affordances*, that consists of a myriad of possibilities for action afforded by the environment (Rietveld & Kiverstein, 2014). Not all of these affordances will invite behaviour equally. Which affordances are more inviting will depend on a host of factors. Whether something is perceived as liftable will for example depend on social setting (a statue in a museum does not afford lifting) and the degree to which a person feels fatigued. Even things such as mental health, for example, have an important role to play: for a depressed person, much less of the environment will be perceived as liftable (De Haan, Rietveld, Stokhof, & Denys, 2013).

1.4. Ecological-enactive learning as the education of attention

On the ecological-enactive approach, the goal of cognitive processes is not truthful or accurate representation, but perceptually guided action. The notion of representation is replaced with *attunement* (Gibson, 1966). How does a child attune to its environment? The ecological answer is through the *education of attention* (Adolph & Kretch, 2015; Gibson, 1953; Gibson, 1963; Gibson, 1966). Attention is understood as selective openness[5] to the field of affordances in relation to a task or goal (Gibson & Rader, 1979; Rietveld & Kiverstein 2014). Attention can only be evaluated with respect to a particular task or goal. In line with the definition of cognition as

---

[4] A good example of this is the explanation Zhu & Bingham (2010) give for the the size-weight illusion. This illusion consists in the fact that a smaller object that has the same mass as a larger one is incorrectly judged to be heavier. However, the illusion disappears if we take people to perceive the 'throw-ability' of the objects. A larger object has to have a greater mass to be just as suitable to throwing as a smaller object. Zhu and Bingham found that the larger object is indeed perceived to be just as heavy as the smaller object when it is equally 'throw-able'.

[5] See Bruineberg and Rietveld (2014) for an ecological-enactive account of the neural dynamics underlying this selective openness.

perceptually guided action, goal-directed behaviour is enabled by a person's sensorimotor skills for acting on relevant affordances. The education of attention thus consists in the development of skills that allow for the context-dependent perceptual guidance of goal-directed behaviour. In other words, as a person is attuned to her environment, she thereby learns how to act on those affordances that are relevant.

In the case of humans, the education of attention mostly takes place in social interaction through processes of *guided rediscovery* (Ingold, 2001; Zukow-Goldring & Ferko, 1994). This means that a master of a practice *shows* or *makes* present affordances to a beginner in such a way that the beginner can learn a particular skill or activity. Note that this process need not be explicitly didactic, but can simply consist in doing something together (Reed, 1995). In child development, the education of attention highlights the active role of both caregiver and child. The caregiver is responsible for setting up the physical conditions that allow the child to learn.[6] Crucial here is the Vygotskian idea of the *zone of proximal development*, which comprises activities that the child can engage in only with the help of the caregiver (Reed, 1995). Besides physical structuring, the education of attention also relies on normative and cognitive structuring of the situation by the caregiver (Williams, 2010). That is, the caregiver's competence in her community's practices allow her to determine what the child is, could, or should be doing, and encourage and discourage the child's behaviour on this basis.[7]

These conditions also hold for the development of more intellectual skills. In learning to talk, a child's utterances do not start out as complete speech acts. Instead, the caregiver treats the child as if she is already competent and on this basis encourages and discourages the child's behaviour.[8] For example, a child uttering something vaguely like /ˈteɪbəl/ in the presence of a table can be encouraged by a caregiver by saying *Very good! That is indeed a table!*, even though nothing in the behaviour of the child shows that she intended to say of is thing that it is a table. Treating the child *as if* she is making an assertion is crucial here. If we were to take

---

[6] For example, Zukow-Goldring and Ferko (1994, p. 172–173) describe a caregiver teaching a child how to roll a ball. The caregiver sits the child down with legs apart, a posture that provides a guide for the trajectory of the ball whilst at the same time allowing the child freedom of movement from a stable position. The caregiver then begins the activity by gathering attention through verbal means (*Look!*), accompanied by moving the ball in the child's line of sight. After the child's attention has been attracted, the caregiver rolls the ball over to the child, and helps the child to roll the ball back, for instance by placing the infants hands on the ball and guiding the return. In this way, the child can gradually learn the activity of rolling a ball.

[7] For example, Rączaszek-Leonardi, Nomi, and Rohlfing (2013) describe how a six-month infant that is being changed accidentally touches a diaper. The mother responds to this contingency by placing the diaper in the child's hand while saying 'yes you can start by holding the diaper'. In this way, the infant's essentially random arm movement is incorporated in the goal directed activity.

[8] Goldstein, King & West (2003) show that 8 month old children are already highly sensitive to positive responses to their vocalisations, a sensitivity that goes beyond mere imitation.

the child to be engaged in mere meaningless babbling, there would be no point at all in encouraging or discouraging her behaviour.

In line with the intimate intertwining of perception and action, successfully participating in her community's practices requires the child to see her world in the correct way (Gallagher, 2005). She has to be able to pick up on those affordances that are salient with regard to the practices of her community. Through the education of attention, the teacher gradually introduces the child in the practices of her community. Educated attention is manifest in altered selective openness: a person is now able to pick up on and skilfully act on affordances of her environment that earlier did not show up for her.

1.5. What a child learns as she learns to speak

On the code view, language learning consists in forming the correct associations between words and mental representations.[9] Ryle (1945a) made a distinction between knowing-that and knowing-how, roughly the distinction between knowing a proposition and having a skill. The code-view assumes linguistic knowledge to be a species of *knowing-that*: knowing a word is knowing *that* it stands for (a set of) mental representations. This does not entail that the code view implies that learning language can have no effect on how the world is represented. It does entail that the child has to mentally represent the world *before* she learns language. On an ecological-enactive approach, this need not be the case. Emphasising the normative structure of the education of attention means that we do not have to assume that the child can already perceive in a categorical way. Instead, by means of the education of attention, the child learns how to 'use words' and *thereby* learns to perceive categorically. In keeping with the relational perspective of the ecological-enactive approach, categorical perception does not rely on mentally representing categories, but consists instead in having the skills for making distinctions between things in the world in line with a community's standards.

This is in line with the observation that a child learns to communicate in 'the midst of "doing"' (Bruner, 1990, p. 70). A child is not a passive observer, but participates *in* activities and it is in these activities that she starts 'doing things with words' (Rączaszek-Leonardi, 2009, p. 170). Viewed from this perspective, the language learner does not acquire knowledge-that *about* language (Taylor, 2013, p. 317). She does not need to learn that a particular words stands for particular things in the world. Instead she comes to *know-how* to do things by talking (Van den Herik, 2017). The question of the acquisition of a language is substituted for asking 'how the child comes to enter

---

[9] Papafragou (2005, p. 357), for example, states that 'the main task of the [language] learner is to figure out which aspects of the input language correspond to which nonlinguistic conceptual primitive notions – or combinations thereof'. Another example, that shows clearly the idea that meanings come before words, is given by Clark (2009, p. 8): 'One issue for language acquisition is how children find out which meanings there are words for; another is just how they map each meaning to the right word.'

the linguistic community' (Reed, 1995, p. 2) which she does by learning a repertoire of social skills (Reed, 1996; Verbrugge, 1985).

This reconceptualisation of language learning has two important implications. First, it shows that linguistic behaviour is always *time-bound* and *context-bound* (Love 1990). According to what Harris (1998, p. 81) calls the *principle of co-temporality*, what is being said is immediately relevant to the current context, and makes sense or fails to make sense within that particular context. In other words, successful linguistic communication consists in the ability of two speakers to converge on a shared understanding (Cuffari, Di Paolo, De Jaegher, 2015; Davidson, 1986), an achievement that relies on bringing all kinds of knowledge to bear (Van den Herik, 2017). Second, language 'must be based on social coordination' (Cowley, 2011b, p. 11). This is evident in ontogenetic development of linguistic abilities: children learn language in a triadic relation, with a caregiver jointly attending to things, in which the focus is on what one can do with language. Language is therefore essentially *dialogical* (Cowley, 2011a). For example, prior to learning its first words, infants already participate in interactions that involve turn-taking behaviours under the guidance of a caregiver (Rączaszek-Leonardi, 2016).

1.6. Utterances of concrete words as attentional actions

To recapitulate, according to the ecological-enactive approach, cognition is a loopy, dynamic, and relational affair. Through the education of attention the developing child gradually attunes to her environment, a process that relies constitutively on both physical and normative structuring of the learning situation by a caregiver that is already competent in the community's practices. However, this leaves open the question what we do with words.

The fact that the ecological-enactive approach eschews mental representations means that the code-view will not do. I start from a different picture of language: *the constraint view* (Rączaszek-Leonardi, 2011, 2016; Rączaszek-Leonardi & Kelso, 2008; Verbrugge, 1985). On the constraint view, language is viewed as a system of social actions that function by constraining unfolding cognitive and interactive dynamics. Because of this reliance on unfolding dynamics, linguistic actions neither have to stand for the processes they constrain, nor do they have to map like a code to the results of their constraining action. In line with the action-oriented nature of cognition, language is reconceptualised as a mode of action. Conceiving of language in terms of constraints brings together the two implications just discussed: ongoing processes are constrained by utterances in real time (language is time-bound and context-bound), and constraints presuppose a communication situation on which the constraining action can take place (the social and dialogical nature of language).

The education of attention model of learning can explain how language gets to have these constraining effects. Language can fulfil the constraining

role because it consists of recognisable and repeatable forms of behaviour, which I refer to as *repeatables* (Van den Herik, 2017). Within the education of attention, the repeatables we construe as words 'anchor' attention[10]: by being easily recognisable, they provide a fixed point that allows for the stabilisation, or sedimentation, of the education of attention.

Different kinds of linguistic actions have different kinds of constraining effect. For the purposes of this paper, I further look into the constraining effects of concrete words from the perspective of ontogenetic development. I start from concrete words because they are a paradigmatic example for proponents of the *code-view*, and empirical research on the relation between language and cognition, which I will consider in the next section, usually involves only concrete words.

The ecological-enactive alternative I propose, however, does not start from 'concrete words' in the sense that code views do. It starts from linguistic actions, or utterances. From the perspective of the child there are no words, there are only repeatables: forms of behaviour that can be *re*cognised from one instance to another, and that can be reenacted. In linguistic behaviour, concrete words usually do not occur in isolation, but rather in the context of a grammatically structured utterance. Notable exceptions are the naming games caregivers play with children, where they point to an object and utter the word for that object, and the child's first utterances. A child's production of language starts out with single-word utterances before she starts combining words into structured sequences (Clark, 2009, p. 115; Fenson et al., 1994).

For the purposes of this paper I start by looking at a child's single word utterances of concrete words. The hypothesis I propose is that a child's first utterances of concrete words are attentional actions. An attentional action is defined as a repeatable performed by a person to *indicate* a particular aspect of the current situation *to* someone *in order to* achieve something. In this paper, I only claim that we can understand a particular kind of linguistic utterance, namely those featuring concrete words, as attentional actions. I do not mean to say that only those utterances are attentional actions, or even that only utterances can be explained as attentional actions. It seems defendable to say, for example, that ostensive gestures are also attentional actions. For the purposes of this paper, however, I will not pursue the question how generalisable the account of attentional actions is.

Let us unpack this definition. It builds on the notion of indication developed by Reed (1995). The child's utterance is not a label, or a name, but a skilful way of directing ecological attention. In other words, we can understand an attentional action as a constraint on attention. While Reed limits the indicational phase to roughly the second year of a child's development, I argue that this account of concrete words as attentional actions can be

---

[10] I take this image picture of words anchoring attention from Clark's (1996) reference to Jackendoff (1996). Given the ecological definition of attention in play in this paper I put it to different use.

generalised to adults' grammatically structured utterances. I return to this issue in §4.2. In first instance, indication is a second-personal skill. We direct somebody else's attention, and whether the attentional action will succeed depends as much on the person that performs it as the person to which it is addressed. Derived from the primordial second-personal action, the person that produces the attentional action can coincide with the addressee, as when we talk to ourselves to direct our own attention.

Note that the definition states that only aspects of the current situation can be indicated. Crucially, the current situation is not limited to what is currently perceivable. A toy that is put away in the closet can be indicated in the same way as a toy that is in view. For the ecological-enactive approach 'there is no sharp line to be drawn between that which is and that which is not perceptually present' (Noë, 2012, pp. 25–29). In general, non-representational approaches such as the ecological-enactive approach often face the objection that they cannot deal with 'representation-hungry' phenomena, such as cognition involving *absentia* or *abstracta* (Clark and Toribio, 1994). See Kiverstein and Rietveld (2018) and Degenaar and Myin (2014) for ecological-enactive solutions to this objection. It is outside the scope of this paper to go into these solutions. For now it is important to see that whereas a code-view conceives of language learning in terms of learning *the* meaning of a decontextualised word, from an ecological-enactive perspective the situated nature of linguistic actions is foregrounded.

There are four elements involved in an attentional action: the person that performs the action, the person that is addressed, the indicated aspect of the situation, and the goal of the attentional action. Attentional actions can be metaphorically 'anchored' to certain aspects of the situation in the unfolding of the education of the attention. In §3 I look at the cognitive mechanism that enable this 'anchoring' in more detail. For now, it is important to note that attention is understood in ecological terms, that is, as the selective openness to the field of affordances in relation to a task or goal. To direct attention by means of an attentional action is thus to foreground an affordance of the current situation, some possibility for action. By thus promoting the salience of certain affordances present in the current situation, the chance that they will be acted upon in the unfolding of the situation is increased. These skills, required for constraining attention and having one's attention constrained, are brought to fruition in the education of attention, which occurs in the context of caregiver and child doing something together.

Realising that attentional actions stem from, and are performed as part of recurrent and structured social situations is crucial for dealing with Quine's question (Reed, 1995): how does a child learn to refer to the same things that their caregivers refer to? Quine (1960) identified the fundamental problem of the indeterminacy of reference. This indeterminacy of reference spells problems for a code-view account of language learning. The argument

is as follows. The code view assumes that a child has to learn what words refer to, which is accomplished by associating the word with the correct mental representation. Adults teach a child a word by pointing out an object and saying the word. This ostensive gesture, however, is not enough to specify what is referred to. The reason for this is that pointing out a dog is indistinguishable from pointing out this particular dog, or a standing-rather-than-a-sitting dog, or all and sundry undetached parts of a dog, or any other fanciful fictions. There is thus a referential ambiguity inherent in ostensive gestures. This means that the child has insufficient evidence for determining what the referent of the ostensive gesture is, and therefore cannot decide on the basis of the ostensive gesture alone what mental representation needs to be associated with the word 'dog'.

In order to be able to answer Quine's question, code-view theorists usually assume that there are a priori constraints on the hypotheses that children entertain with respect to word-meanings (Clark, 2009, p. 124). For example, children are thought to assume that words refer to a whole objects (*the whole-object assumption*) and taxonomic kinds, rather than thematically related objects (*the taxonomic assumption*; Clark, 1981, p. 40; Markman, 1981). By means of these *a priori* constraints, the referential ambiguity is dissolved and a child is able to arrive at the correct hypothesis, for example that 'dog' refers to dogs.

If we, however, assume that utterances of concrete words are attentional actions, Quine's question can be evaded. To understand why this is the case, we have to realise that the environment of the child is highly structured, and consists of recurring situations, routines, and activities. It is in these activities that a child starts responding to and performing attentional actions. The task that the child is faced with is becoming an active participant in the activities she engages in with her caregivers. Unsurprisingly, the first twenty words that a child produces typically pertain to salient aspects of recurrent situations that are highly relevant to the infant, such as people (*daddy, mommy*), games and routines (*bye, night-night*), animals (*dog, kitty*), and toys (*book, balloon*) (Fenson et al., 1994). Of these twenty first words, fourteen are concrete words.[11]

An interesting observation is made by Clark (2009, p. 88), who notes that when children start using language, they use a lot of deictic terms like *that* and general purpose verbs like *do*. As she points out, 'without contextual details … it is usually impossible to interpret such utterances', whereas, in 'context, with the aid of joint attention, it is normally quite clear what children are talking about when they do this.' The crucial insight is that this line of reasoning holds for all of the child's initial linguistic behaviour, and continues to hold, although to a lesser degree, even for fully competent adult linguistic

---

[11] The concrete words are *daddy, mommy, dog, ball, baby, book, shoe, kitty, bird, eye, balloon, bottle, banana*, and *juice*, the other words are *bye, hi, uh oh, no, woof*, and *night night* (Fenson et al. 1994, Table 15, p. 93).

behaviour.[12] This is because attentional actions function similar to ostensive pointing gestures: they are performed in a context in order to foreground a particular aspect of the situation, and without contextual details, they too are impossible to interpret. Quine's question spells trouble only if we take the ostensive gesture as it is used in the education of attention as evidence for determining the reference. If we instead understand the attentional action to, as it were, take the place of the ostensive gesture, we do not have to answer Quine's question.

In other words, the attentional action inherits the referential ambiguity of the ostensive gesture by means of which the education of attention is initially achieved. This is possible because an attentional action is a species of ostensive gesture. As Baggs (2015, p. 260) puts it, speaking is 'a technique for pointing'. The referential ambiguity only emerges from an observer-perspective, once an observer asks himself what *the* reference of *this* word is. Learning to perform an attentional action does not require taking a stand on what a word refers to. Determining the extension of an attentional action (that which can be indicated by performing the attentional action) would be like determining the extension of the pointing gesture abstracted from the context in which the pointing gesture occurs: a pointless exercise.

Crucially, the referential ambiguity usually does not show up from a participant's perspective, that is, from the perspective of the child or the caregiver in the education of attention, and is not detrimental to communicational success. The reason is that the object of indication is individuated, not by decontextualised referential knowledge (of the form *'dog' refers to dogs*), but by the unfolding of the (recurrent) situation, in which only certain affordances are relevant, and thus can become more salient as a result of the attentional action.

At the same time, to highlight the situatedness and ostensive nature of indication by means of attentional actions is not to say that there is no potential for misunderstanding or conflict. It is crucial to note here that *indicate* is used as a 'verb of success' (Ryle, 1949/2009, p. 114). As mentioned, whether someone succeeds in indicating something to someone is not only dependent on her behaviour, but lies in the successful coordination of behaviour. In other words, 'indication of a single object, place, or event, is an *achievement* of a dyad, not a given.' (Reed, 1995, p. 11). In order to play her part in this achievement, a child needs to be able to successfully coordinate her behaviour with others by means of attentional actions. This

---

[12] For an extension to competent adult language use, we can remind ourselves of Ryle's (1945b, p. 215) claim that all words incorporate a systematic ambiguity: 'A given word will, in different sorts of context, express ideas of an indefinite range of differing logical types and, therefore, with different logical powers.' The example he gives is that of punctual, which 'can be used to characterize a person's arrival at a place, the person who arrives there, his character and even the average character of a class of persons'. Comparing these different kinds of punctuality would be absurd, according to Ryle, and thus he concludes, 'where precision is wanted, it is wrong to speak of "the idea" of punctuality' (Ibid., 216). This conclusion is in line with our rejection of the code-view.

means that the success conditions are not specified by the attentional action, but emerge in the unfolding of the situation and are to be decided upon by the interacting dyad.

To summarise, on the constraint view, concrete words do not stand for anything. Learning to use these words does not consist in acquiring knowledge-that about their meaning. Instead, learning to talk is learning a repertoire of social skills. A child's first utterances of concrete words are used to indicate aspects of the current situation to others in order to achieve something, and are therefore understood as *attentional actions*. An attentional action is like an ostensive gesture: it points out something. Learning takes place by means of the education of attention, where attentional actions are 'anchored' to aspects of the environment. In order to understand better how this 'anchoring' functions, we now turn to empirical research into the effect of colour words on categorical colour perception.

2. The case of colour

Categorical colour perception has been the preferential paradigm for studying the influence of language on putatively non-linguistic processes, because it is a prime example of categorical perception that allows for easy cross-cultural comparison. In particular, research has focussed on category effects, which can broadly be understood as any effect of verbal categorisation on colour cognition; for example, an effect of verbal categorisation on speed or accuracy in discriminating or remembering colours. In this section, I report the *category effects* found in this research, in order to draw some implications for understanding utterances of concrete words as attentional actions in the next section.

2.1. Basic colour words and cultural differences

Research into categorical colour research relies on the notion of basic colour words. Colour words are basic when they are monolexemic, not subsumed under other colour words (such as *navy*), applicable to all objects (which excludes *blonde* for example), and psychologically salient for a community (Berlin & Kay, 1969). In English, these criteria are met by the words *white, black, red, green, yellow, blue, brown, purple, pink, orange*, and *grey*. Other languages, however, employ different colour words that categorise colour space in different ways.[13] For example, the Dani from New Guinea employ just two basic colour words, *mili* and *mola* (Rosch-Heider & Olivier, 1972), and

---

[13] There is a debate whether the development of colour terms across linguistic communities is universal. Berlin and Kay (1969) for instance, argued that all linguistic communities go through the same seven stages, starting out by making a dark-cool/light-warm distinction, then adding red, and so forth until they finally make it to the universal end-point which is expressed in the eleven English basic colour words. There however, seem to be counter examples to this alleged universality (Roberson, Davies, & Davidoff, 2000). Moreover, the alleged universal order is not found in language learning (Andrick & Tager-Flusberg, 1986; Pitchford & Mullen, 2002).

the Berinmo have five basic colour words including *nol* which spans colours ranging from *yellow* to *purple* (Roberson, Davies, & Davidoff, 2000).

2.2. Russian blues

To investigate category effects, Winawer et al. (2007) devised a speeded colour discrimination task that relies on the fact that Russian speakers have basic colour words for lighter blues (*goluboy*) and darker blues (*siniy*), where the English basic colour word *blue* spans these two categories. Participants were shown a triad of blue colour squares and had to indicate as quickly and as accurately as possible which of the bottom two squares was the same colour as the top square (see Figure 2). In so-called cross-category trials, the distractor square was across the *siniy/goluboy* category border, as determined for that each Russian speaker individually by means of an elicitation procedure, whereas in within-category trials the distractor square was in the same verbal category as the target square. Of course, for English speakers all squares always belonged to the same basic verbal colour category, *blue*.
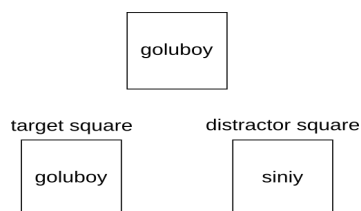

Figure 2. Example of a cross-category trial in Winawer et al. (2007).

The results of this speeded discrimination task were that Russian speakers' performance showed a category effect: their reaction times were significantly shorter in cross-category trials than in within-category trials, an effect that was absent for English speakers. In other words, when the target and distractor square fall on different sides of the *siniy-goluboy* category boundary, Russian speakers are able to identify the target square faster than when both the target and and the distractor square fall in the same category. Interestingly, when asked to differentiate between *light blue* and *dark blue*, English participants drew almost the same border as Russians did between *goluboy* and *siniy*. This shows that English participants can visually distinguish light and dark blues in similar ways to Russian participants.

This experiment evaded the critique levelled against earlier research. For example, Roberson, Davidoff, Davies, and Shapiro (2005) asked participants to categorise a collection of colour swatches, and found that people tended to categorise the colour chips in accordance with verbal categories. They interpreted the results as showing a category effect: they claimed that it showed that to the participants, colours from the same verbal category looked

more alike than colours from a different verbal category. However, this effect can also be explained in terms of a deliberate linguistic strategy: participants could simply put together those colours for which they use the same colour word (Pinker, 1994). In the experiment of Winawer et al. (2007), this strategy was not possible: the task did not involve subjective judgements of colour similarity, but consisted in finding the correct answer to a simple discrimination task. Moreover, the effect was measured in the reaction time of participants, a measure that is taken not to be under the participants' control.

Further evidence that category effects are not due to a deliberate linguistic strategy is given by Thierry et al. (2009). They devised an experiment that showed that the category boundary in Greek between basic colour words for light and dark blue has a category effect on implicit colour discrimination in a colour oddball detection task. In this experiment, participants were shown a sequence of stimuli differing in colour and shape and were instructed to press a button when they saw a particular shape. This means that changes in colour were irrelevant to the task. However, Thierry et al. were able to find a category effect using an electrophysiological measure: the Greek distinction between light and dark blues led to a larger visual mismatch negativity ('an index of automatic and preattentive change detection' Ibid., p. 4567) in the case of a deviant colour stimulus as compared to English participants for whom the deviant stimuli fell in the same verbal category (cf. Athanasopoulos et al., 2009).

The category effects just discussed could be explained in two different ways. First, the effect might be due to 'warping' of perceptual space on longer timescales. On this explanation, development in a culture with specific verbal colour categories leads to lasting effects on colour perception. Second, the effect might unfold on much shorter time-scales. On this explanation, linguistic processes have an online modulatory effect on cognitive processes. A way to decide between these two competing explanations is to compare the interference of verbal dual tasks and other cognitively demanding duals tasks on the category effects. If the category effects obtain due to warping of perceptual space on longer timescales, a verbal dual task should not lead to more interference when compared to a equally demanding non-verbal dual task. If, however, the effect is due to online modulation a verbal dual task should lead to interference over and above that of non-verbal dual tasks.

Earlier studies found that the category effects were indeed subject to verbal interference in this way. For example, Roberson and Davidoff (2000) found that the category effect that obtained in a two-alternative forced-choice recognition experiment[14] could be eliminated if participants were asked to recognise the colour swatch they had just seen after reading aloud for five to ten seconds (cf. Pilling & Davies, 2004). This result only shows verbal

---

[14] This means that participants were shown a colour swatch, and after a delay of five or ten seconds were shown a target and distractor square, and had to indicate which of these was the same as the initial colour swatch.

interference in the case of colour memory. Winawer et al. (2007) also found that their observed category effect was eliminated by a verbal-interference dual task (silent rehearsal of digit strings), but not by a spatial-interference dual task (memorising a spatial pattern). The elimination of the category effect is thus not due to some general cognitive load limitation. However, verbal interference does not eliminate all category effects (Pilling, Wiggett, Özgen, & Davies 2003), a finding that suggests that category effects also play out on longer timescales.

The linguistic origins of the category effects are further corroborated by findings showing that they are strongest in the right visual hemifield, which projects contralaterally to the putative language-dominant left hemisphere (Drivonikou et al., 2007; Franklin, Drivonikou, Bevis, et al., 2008; Franklin, Drivonikou, Clifford et al., 2008; Gilbert, Regier, Kay, & Ivry, 2006; Regier & Kay, 2009; but see Suegami, Aminihajibashi, & Laeng, 2014).

2.3. Categorising and anomic aphasia

Further evidence of the transient effect of language on perception comes from studying individuals with *anomic aphasia*, that is, who have difficulty naming, among other things, colours. Here the influence of linguistic categories on putatively non-linguistic tasks is shown to be more dramatic: without an active command of colour words the task of colour categorisation is impossible (Davidoff & Luzzatti, 2005; Dummett, 1975). Davidoff and Roberson (2004) describe the problems that patient LEW has with sorting colour chips into categories. Instead of relying on categorical perception, he had to rely on discrimination:

> if colour samples were presented for which the within-group similarity was much greater than the between-group similarity (i.e., narrow ranges of reds, greens, yellows, and blues), he sorted them into four groups without error. Nevertheless, his performance was abnormal because he used a slow pairwise comparison for each stimulus; the colour groups did not 'pop-out'. His abject failure was for tasks where within-group colours had a wide range of lightness and saturation; in those situations, assessing visual similarity is extremely difficult. (Davidoff & Roberson, 2004, p. 139).

In the first task, the greater within-group similarity allowed LEW to arrive at the grouping using only the ability to perceptually discriminate between two colour samples when they were present at the same time. When the within group similarity was lower, LEW was presented with a problem that is unsolvable through discrimination abilities alone. This closely resembles the *sorites* paradox (Davidoff & Roberson, 2004): we can make a row of colour chips ranging from red to blue, and make sure that the difference in hue between all the chips is equal. By relying only on colour discrimination, there is no point discernible where the category boundary should be. And indeed,

when LEW was confronted with such a task, he drew the conclusion that all colour chips belonged to the same colour category. He arrived at this conclusion by looking for colour chips that appeared identical to him, and then moving them to a group, and then repeating this procedure until all the colour chips were in the same group.

2.4. Categorical perception beyond colour

Categorical colour perception is usually taken as a paradigm for studying the effects that verbal categories have on putatively non-verbal cognitive processes. However, if the category effects were only obtained in categorical colour perception, they would at best provide a cognitive curiosity. Fortunately, there is a growing literature which aims to extend the findings from the research on categorical colour perception to categorical perception in general (see Lupyan, 2012, for an overview). Gilbert, Regier, Kay, and Ivry (2008), for example, found that the the lateralisation of the category effect also occurs for the perception of animal figures (cats and dogs). This asymmetry between hemifields was, like the category effect in colour perception, subject to interference in the case of verbal dual task, but not a spatial dual task. A second example is the category effect that Boutonnet, Dering, Viñas-Guasch, and Thierry (2013) found using electrophysiological measures for the categorical distinction English speakers draw between cups and mugs, a distinction that Spanish speakers do not make. A final example is an experiment performed by Lupyan and Spivey (2010), where participants had to attend selectively to four fives presented together with four 2s while fixating their gaze on the cross (see Figure 3). As soon as a small dot appears next to one of the 5s, the subjects have to press a button. The subjects were found to perform faster on trials where they heard a recorded voice saying 'attend to the five' as compared to trials where they heard a recorded voice saying 'attend to the category'. As the task of attending to the 5s remained the same for a 45-minute period of time, hearing the word 'five' was completely redundant. This result seems to suggest that hearing the word 'five' directed the participants' attention to the 5s.

```
        2       5
    2           5
          +
    5           2
        5       2
```

Figure 3. Display used by Lupyan and Spivey (2010, based on their figure 1).

3. Attentional actions as constraints on phenotypic reorganisation

In this section, I combine the account of utterances of concrete words as attentional actions based on the constraint view with the empirical results discussed in the previous section. In line with the *don't-need strategy* as discussed in §1.2, I provide an account that can potentially explain the effects suggested by the empirical research without invoking representations.

The empirical results suggest that there are indeed category effects of verbal categories on colour perception. At the same time, the fragility of these category effects, as manifested in their being subject to verbal interference, suggests that these category effects are the result of a complex interplay of different timescales. The empirical evidence also shows that the category effects based in verbal categorisation are found in the absence of overt language use. A potential way of understanding the mechanism behind this is the ecological notion of a *task-specific device* (Bingham, 1988; Runeson, 1977; Wilson & Golonka, 2013). The guiding idea is that for the purposes of a particular task, an organism self-organises into a task-specific device which is assembled of resources distributed over body, brain, and environment.[15]

One way of understanding the assembly of task-specific devices is by means of *phenotypic reorganisation*. On this view, 'people turn themselves into walkers, throwers, graspers, and so on, and in so doing, they perceive the world in relation to what they have become.' (Proffitt & Linkenauger, 2013, p. 172). As can be seen in this quote, Proffitt and Linkenauger focus on basic sensorimotor interactions with the environment. Their claim is that visual ecological information for affordances[16] is scaled based on morphology, physiology, and target-directed action. For example, whether something is graspable for a particular person depends, amongst other variables, on the size of their hand. Therefore, when a person is engaged in *grasping* something, the information for *graspable* is scaled based on hand size, among other things. In this way, the purposes of the organism, in the sense of the activities the organism is currently engaged in, 'mandate a goal-directed phenotypic reorganization' (Ibid., p. 180). Note that the ongoing phenotypic reorganisation is envisaged as a dynamic process of self-organisation.

For our present concerns, we can extend the notion of becoming a task-specific device by means of phenotypic reorganisation to categorical

---

[15]   For a discussion of the role the brain may play in this process of self-organisation and how neuroscience might develop methods to study this, see Anderson (2014, pp. 272–280) and Van Elk, Slors, and Bekkering (2010). See Bruineberg, Kiverstein, and Rietveld (2016) for an ecological-enactive interpretation of the free energy principle that similarly describes the self-organisation of the brain-organism-environment system.

[16]   Ecological information is information without content, that is non-semantic information, that enables an organism to pick up on affordances. The basic idea is that structures in ambient energy arrays are informative to an organism because they specify certain aspects of the environment. In other words, ecological information is information *for* affordances, not information *about* the world. Ecological information is not transmitted, nor is it processed or stored by a cognitive system (Gibson 1979). For a discussion see Van Dijk, Withagen, and Bongers (2015).

colour perception by taking the social context into account. Information then is scaled based not on characteristics of the individual, such as hand size, or actions performed individually, but rather based on community level patterns of categorisation as inculcated through the education of attention and the (attentional) actions of others. As we noticed earlier, the normative structuring of the education of attention entails that the child need not already be able to perceive categorically in order to be taught to do so.[17] What task a child is engaged in is determined initially by means of the normative and physical structuring of the learning situation by someone who has already mastered the relevant practice. If a caregiver introduces the child to colour words, she educates the child's attention by structuring the learning situation. For example, by drawing the child's attention to coloured objects, by showing the sorting of coloured objects, and so on. The caregiver also provides feedback on the child's behaviour. In this process, the child attunes to the culturally determined colour categories. The use of colour words as attentional actions is crucial here, for they provide the caregiver with the means to 'calibrate' the child. Hearing the word 'red' thereby comes to acts as a constraint on ongoing phenotypic reorganisation, making the child into a better red detector.

Gibson (1966, p. 52) defined the education of attention as 'a greater noticing of the critical differences with less noticing of irrelevancies' and 'a progressive focussing or centering of the perceptual system'. The current account suggests that this progressive focussing needs to be understood in a dynamic way: as a person interacts with his or her environment, there is a continuous phenotypic self-reorganisation into task specific devices. This phenotypic reorganisation is constrained by the current situation, which in the human case often includes attentional actions. But on this account, attentional actions are not the only constraints that shape phenotypic reorganisation in culturally specific ways. The example of categorical colour perception suggests that attending to something that has a particular colour can act as a constraint on the ongoing phenotypic reorganisation. Attending to something red can have a similar constraining effect as hearing the colour word 'red', where both lead to phenotypic reorganisation as a red detector. In other words, and in line with the loopy nature of cognition, the ongoing phenotypic reorganisation is perceptually guided.

On the level of neural dynamics, the effect of verbal interference can be explained by adopting a neural reuse perspective. Anderson (2016, p. 1) defines neural reuse as a form of neuroplasticity that entails that a 'diverse behavioral repertoire is achieved by means of the creation of multiple, nested, and overlapping neural coalitions, in which each neural element is

---

[17] This of course does not entail that the child requires no abilities whatsoever. For example, learning co-lour words requires that the child is able to discriminate between different coloured objects. But colour discrimination does not necessarily rely on categorical colour perception, as the example of LEW discussed in §2.3 shows.

a member of multiple different coalitions and cooperates with a different set of partners at different times.' A verbal interference task might rely on neural resources that would otherwise be used in the phenotypic reorganisation as a categorical colour perceiver. As Lupyan (2012, p. 4) remarks, once we have acquired categorical perception, much of our experience might best be described as 'hybrid visuo-linguistic experience.'

The current account has no direct implications for phenomenal experience. One possibility is given by Anderson (2014, p. 271), who states that 'Language probably does not make the world look different, but it can make some things easier to see'. Another possibility is that language does make the world look different, and thereby makes some things easier to see. For the purposes of this paper I will not explore the different ways of drawing the phenomenological implications of the current account.

In line with the idea that organisms are always selectively open to a field of affordances (Rietveld & Kiverstein, 2014), any organism self-organises into many different task specific devices at any given time. In line with the neural reuse perspective, we should be careful not to conceive of these task specific devices as neural modules, but rather as dynamically self-organising assemblies of resources distributed over body, brain, and environment. For example, a person is never only a grasper, although the anticipation of the act of grasping will lead to a shift in attention, which can be understood as a phenotypic reorganisation that enables an enhanced selective openness to those aspects of the environment that are relevant for grasping. The phenotypic reorganisation into task-specific devices is thus not an all or nothing affair. It is not the case that hearing the word 'red' either results in the phenotypic reorganisation into a red-detector, or fails to do so. Instead, we can think of attentional actions as transiently modulating phenotypic reorganisation. This modulation can then be up-regulated, for example by means of attentional actions, or down-regulated, for example by means of a verbal interference task (cf. Lupyan 2012).

This account of task specific devices, when combined with neural reuse, shows how perception can be constrained by processes at multiple time scales. The resulting picture is the following: by means of the education of attention, individuals learn to perceive categorically as expressed in structures in the dynamic unfolding of phenotypic reorganisation. This phenotypic reorganisation is perceptually guided and constrained by the actions of oneself as well of those of others. One example of these constraining actions are what I have called attentional actions. Although I have taken colour as a case study, the results for other forms of categorical perception discussed in the previous section suggests that this account might be applicable to categorical perception at large. For example, hearing the word 'chair' makes the person into a better 'chair detector' (Lupyan & Swingley, 2012).

4. Language beyond attentional actions

So far, I have argued that utterances of concrete words are best taken to be attentional actions (§1). I then discussed research on the effect of community colour categories on individual colour cognition (§2), and proposed an ecological mechanism based on the idea that attentional actions function as constraints on the dynamic unfolding of phenotypic reorganisation (§3). In this final section, I situate the account of utterances of concrete words as attentional actions in a larger theoretical perspective on language by means of three remarks on (1) the context-dependency and reflexivity of language, (2) an extension of the current account to grammatically structured language, and (3) written language.

4.1. Context-dependency and reflexivity

It is important to stress that the current account highlights the context-dependency of linguistic actions. Attentional actions are not *stored* and *retrieved*, but *produced* to constrain processes that are unfolding *now*. In other words, we do not *use* language, understood as the tokening of culturally determined types, we *make* language in line with a history of learning (Harris, 1980). Although it is certainly the case that we can use words in many *different* contexts, this only shows our ability to recontextualise. And this ability to *re*contextualise does not entail a decontextualised meaning. To be able to walk on many different surfaces does not entail an ability to walk *in abstracto*; similarly, being able to use a word in different or novel contexts does not entail an ability to use that word *in abstracto*, that is, it does not require knowledge of a decontextualised 'meaning'. This point is captured by Voloshinov (1930/1973, p. 68) when he says that 'the task of understanding does not basically amount to recognizing the form used, but rather to understanding it in a particular, concrete context' (cf. Davidson, 1986). Recognising the context-dependent nature of linguistic actions also does not negate the fact that we have normative practices that aim to regularise and standardise linguistic actions across contexts. As we have seen, the normative structuring of the initiate learning situation by the caregiver is critical for the child's education of attention. Other paradigmatic examples include extensive institutionalised language education and the writing of dictionaries and grammar books and the prescriptive use made out of these codifications. These latter examples are an expression of the reflexivity of language, that is, our ability to talk about talking (Harris, 1998; Taylor, 1992, 2000, 2010, 2013). I do not object to taking this reflexivity and the normative enforcing of linguistic behaviour it affords seriously. What I do object to is to take 'products of these [reflexive] processes as *realia*, and to retroject them on to languagers as the basis for their languaging activities' (Love, 2017, p. 1). To understand this, take for example the practice of promise making. This is

a reflexive practice because it consists in taking some linguistic act to be a promise. This is to say, if we were unable to say of some act that it was a promise, and we were unable to determine whether a promise was kept or broken, etc, we could not have the practice of promise making. At the same time, making promises does not have to be grounded in the existence of mental promises: the normative import of making a promise is not guaranteed by a mental state, but instead, by actually making the promise. Similarly, our practices that stabilise and normatively enforce the meanings of words do not have to be grounded in mental meanings. To reify promises or meanings is a prime example of the fallacy of misplaced concreteness (Whitehead, 1929). In line with this focus on context-dependency, the constraining effect of attentional actions will also be context dependent. So, for example, the constraining effects of being asked the reflexive question *how do you spell the word 'green'* will be very different from a person telling you to *look at that green one*. For the purposes of this paper, I have only explored situations which are like the latter, in the sense that they involve what we might call, in line with the notion of indication in the definition of attentional actions, the *indicative use* of a concrete word. The question whether the reflexivity of language can be explained in terms of constraints I leave for a future occasion.

4.2. Grammatical structure and attentional actions

In my discussion on attentional actions, I have mainly considered a child's single-word utterances of concrete words. This leaves open the question how we should make sense of concrete words as they are used in grammatically structured utterances. A tentative proposal is to see such utterances as enabling more fine-tuned attentional actions. As we have seen, attentional actions are performed by someone to indicate an aspect of the situation to someone else in order to achieve something. A child's attentional action *ball* is used to indicate a ball in order, for instance, to request a ball that is already the focus of attention, to initiate a game that child and caregiver regularly play together, to inquire as to the whereabouts of a particular ball that the child lost, to mark the sudden unexpected appearance of a ball, and so forth. In this single word attentional action, only the object of indication is thus explicated.

The tentative proposal is that in structured attentional actions, concrete words still serve this indicative role. This means that the discussion on attentional actions as constraints on phenotypic reorganisation also applies to structured attentional actions. For example, an utterance of the form *could you give me the ball* is akin to a child's single-word utterance 'ball' in the sense that it is an attentional action by means of which a ball is indicated in order to achieve something (in this case the other person handing me the ball). However, whereas only the structured environment reduces potential ambiguity in the case of the child's single word attentional action, this potential ambiguity is reduced by linguistic means in the case of the structured

attentional action. If I'm asking another person *Could you give me the ball*, this is more constraining than if I simply say *ball*; but both are attentional actions that indicate the ball in order to achieve something.

In terms of constraints, a structured attentional action is thus more constraining than a single-word attentional action, but the same basic structure is preserved when the child makes the transition to structured attentional actions. That is, structured attentional actions still function akin to ostensive gestures.[18] Of course, not all structured linguistic actions are attentional actions. However, providing criteria for delineating which structured linguistic actions are attentional actions and which are not falls outside the scope of this paper.

One might object that an extension to structured attentional actions fails the criterion of providing a don't-need account, as linguistic structure requires mental representation. And indeed, the received view of the epistemology of language *qua* linguistic structure is that it requires an explicit representation of the rules that together can be said to constitute the grammar of a language.[19] However, non-representational alternatives have also been proposed. Christiansen and Chater (2015, p. 17), for example, suggest that linguistic structure can be understood in terms of a history of processing operations that constrain current processing. They use a metaphor also used by enactivists of 'laying down a path in walking' (Varela, Thompson, Rosch, 1991). According to this metaphor, we conceive of the unfolding of current processes as being constrained by the traces left by earlier processes. An example is a desire path: when a person walks on the grass, she leaves a visible trail that can guide the next person's walking behaviour. By many people thus walking in each other's footsteps, the path becomes more visible and thus constrains subsequent walking behaviour more strongly. Based on this history of processing account of linguistic structure, we can conceive of linguistic structure as providing further constraints. The resulting image is thus that when a child learns to employ structured utterances, her attentional actions become more strongly constraining. I realise that what I have said here with respect to linguistic structure is very brief, but it is merely intended to provide a direction in which the current account can be extended.

---

[18] Kukla (2017) goes as far as saying that speech acts that make highly theoretical claims still have an ostensive dimension. She bases her account on Heidegger and Haugeland's interpretation (2013, p. 67), who claims that in 'making an assertion a speaker lets what is being talked about show itself from itself, by pointing it out—putting it on exhibit, so to speak. If, for instance, I discreetly mention that your shoelace is untied, I draw the shoelace to your attention so that you can see, "from" the shoelace itself, that it is untied. By pointing out the untied shoelace (something I could also do without words), I let it be seen— let it show itself from itself.'

[19] Matthews (2003, pp. 188–189) defines the 'Received View' of linguistic knowledge as follows: 'knowing a language is a matter of knowing the system of rules and principles that is the grammar for that language. To have such knowledge is to have an explicit internal representation of these rules and principles.' Devitt (2006), similarly, claims that 'the received view' is that 'language processing involves metalinguistic representations of the syntactic and semantic properties of linguistic expressions'.

4.3. Written language

The third remark situating the account of attentional actions concerns the extension to written language. First note that written language contributes considerably to the normative reflexive enforcing of linguistic behaviour. For example, without written language we could not write dictionaries, and we would thus be unable to 'look up' the meaning of a particular word (Ong, 1982). Moreover, for us, hyper-literate human beings, our experience of spoken language is shaped profoundly by our facility with written language (Harvey, 2015). On the code view, a spoken utterance can encode *the same* mental meaning as a written sentence. The constraint view, however, highlights the fact that the production of spatio-temporal patterns in spoken or sign language and the production of spatial patterns in written language belong to different ontological categories, are used to do different things, and constitute different cognitive domains with different cognitive dynamics (Kravchenko, 2007, 2009; Linell, 2005; Love, 2007). Perhaps the most conspicuous difference is that for writing and reading the principle of co-temporality, as discussed in §1.5, does not hold. Understanding written concrete words from a constraint-view thus requires an additional account, that is outside the scope of the current paper.

5. Conclusion

I argued that on an ecological-enactive approach, utterances of concrete words are best seen as attentional actions that are used to indicate situations, events, objects, or characteristics thereof in order to coordinate behaviour. The results from empirical research show that the attention-directing effects of socially constituted categories affect cognitive processes in the absence of overt language use. This can be understood by relying on the idea that humans continuously self-organise into task specifics devices, a form of phenotypic reorganisation. The unfolding of this process is constrained by a history of learning, the actions of oneself, others, as well as aspects of the environment. In this way this paper proposed a novel explanation for the empirical phenomenon of category effects. Finally, I situated the account of concrete words as attentional action in a larger theoretical perspective on language. Although a small step, the account of spoken concrete words as attentional actions shows how the ecological-enactive approach to cognition can be extended to explain linguistic behaviour.

6. Acknowledgements

audience at the 13th Workshop of the OZSW Study Group Enactivism for comments and discussion. The constructive comments of two anonymous reviewers were also very helpful. Work on this paper was financially supported by The Netherlands Organisation for Scientific Research (NWO).

7. References

Adolph, K. E., & Kretch, K. S. (2015). Gibson's theory of perceptual learning. *International Encyclopedia of the Social & Behavioral Sciences, Second Edition*, *2015*, 127–134. doi:10.1016/B978-0-08-097086-8.23096-1

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, *149*(1), 91–130. doi:10.1016/S0004-3702(03)00054-7

Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. Cambridge, MA.: The MIT Press.

Anderson, M. L. (2016). Précis of after phrenology: Neural reuse and the interactive brain. *Behavioral and Brain Sciences*, *39*, 1–45. doi:10.1017/S0140525X15000631

Andrick, G., & Tager-Flusberg, H. (1986). The acquisition of colour terms. *Journal of Child Language*, *13*(1), 119–134. doi:10.1017/S0305000900000337

Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J., & Thierry, G. (2009). The Whorfian mind. Electrophysiological evidence that language shapes perception. *Journal Communicative & Integrative Biology*, *2*(4), 332–334. doi:10.4161/cib.2.4.8400

Aristotle (1975). *Categories and De Interpretatione*. J.L. Ackrill (Trans.). Oxford: Clarendon Press.

Baggs, E. (2015). A radical empiricist theory of speaking: Linguistic meaning without conventions. *Ecological Psychology*, *27*(3), 251–264. doi:10.1080/10407413.2015.1068655

Bakhtin, M. M. (1982). *The dialogical imagination* (C. Emerson & M. Holquist, Trans.). University of Texas Press.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. California, CA: University of California Press.

Bingham, G. P. (1988) Task-specific devices and the perceptual bottleneck. *Human Movement Science*, *2*(2–4), 225–264. doi:10.1016/0167-9457(88)90013-9

Boutonnet, B., Dering, B., Viñas-Guasch, N., & Thierry, G. (2013). Seeing objects through the language glass. *Journal of Cognitive Neuroscience*, *25*(10), 1702–1710. doi:10.1162/jocn_a_00415

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139 – 159.

Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances, *Frontiers in Human Neuroscience*, *8*(599), 1–14. doi:10.3389/fnhum.2014.00599

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2016). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese*, 1–28. doi:10.1007/s11229-016-1239-1

Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Cairns, H. & Cairns, C. (1976) *Psycholinguistics*. New York, NY: Holt, Rinehart & Winston.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: The MIT Press.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioural and Brain Sciences*, 1–27. doi:10.1017/S0140525X1500031X

Clark, A. (1996). Linguistic anchors in the sea of thought? *Pragmatics & Cognition*, *4*(1), 93–103. doi:10.1075/pc.4.1.09cla

Clark, A. (1997). *Being there*. Cambridge, MA.: The MIT Press.

Clark, A., & Toribio, J. (1994). Doing without representing? *Synthese*, *101*, 401–431.

Clark, E. V. (1981). Acquisitional principles in lexical development. In: S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on language and thought interrelations in development* (pp. 31–71). Cambridge: Cambridge University Press.

Clark, E. V. (2009). *First language acquisition. Second edition*. Cambridge: Cambridge University Press.

Cowley, S. J. (2011a). Distributed language. In: Cowley, S. (Ed.), *Distributed language* (pp. 185–210). Amsterdam/Philadelphia, PA: John Benjamins.

Cowley, S. J. (2011b). Taking a language stance. *Ecological Psychology*, *23*, 1–25. doi:10.1080/10407413.2011.591272

Cuffari, E. C., Di Paolo, E., & De Jaegher, H. (2015). From participatory sense-making to language: there and back again. *Phenomenology and the Cognitive Sciences*, *14*(4), 1089–1125. doi:10.1007/s11097-014-9404-9

Cruse, D. A. (2017). The Lexicon. In: M. Aronoff & J. Rees-Miller (Eds.), *The handbook of linguistics 2nd edition* (pp. 235–254). Oxford: Wiley Blackwell.

Davidoff, J., & Luzzatti, C. (2005). Language impairment and colour categories. *Behavioral and Brain Sciences*, *28*(4), 494–495. doi:10.1017/S0140525X05280081/

Davidoff, J., & Roberson, D. (2004) Preserved thematic and impaired taxonomic categorisation: A case study. *Language and Cognitive Processes*, *19*(1), 137–174. doi:10.1080/01690960344000125

Davidson, D., 1986. A nice derangement of epitaphs. In E. Lepore (Ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (pp. 433–446). Oxford: Blackwell.

Davis, H.G., Taylor, T. J. (Eds.). (1990). *Redefining linguistics*. London/New York, NY: Routledge.

Davis, H.G., Taylor, T. J. (Eds.). (2003). *Rethinking linguistics*. London/New York, NY: RoutledgeCurzon.

Degenaar, J., & Myin, E. (2014). Representation hunger reconsidered. *Synthese*, *191*, 3639–3648. doi:10.1007/s11229-014-0484-4

Devitt, M. (2006). *Ignorance of language*. Oxford: Clarendon Press.

De Haan, S., Rietveld, E., Stokhof, M., & Denys, D. (2013) The phenomenology of deep brain stimulation-induced changes in obsessive-compulsive disorder patients: An enactive affordance-based model. *Frontiers in Human Neuroscience*, *7*, 1–14. doi:10.3389/fnhum.2013.00653

De Saussure, F. (2011). *Cours de linguistique generale (W. Baskin, Trans.)*. New York, NY: Columbia University Press. (Original work published in 1922).

Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R. L., (2007). Further evidence that Whorfian effects are stronger in the right visual field than the left. *PNAS*, *104*(3), 1097–1102. doi:10.1073/pnas.0610132104

Dummett, M. (1975). Wang's paradox. *Synthese*, *3*(4), 201–232. doi:10.1007/BF00485048

Engel, K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*, *17*(5), 202–209. doi:10.1016/j.tics.2013.03.006

Fenson, L., Dale, P. S., Reznick, S. J., Bates, E., Thal, D. J ., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*, (5, serial No. 242).

Fodor, J. A., (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. A. (2008). *LOT 2: The language of thought revisited*. Oxford: Oxford University Press.

Franklin, A., Drivonikou, G. V., Clifford, A., Kay, P., Regier, T., Davies, I. R. L. (2008). Lateralization of categorical perception of color changes with color term acquisition. *PNAS*, *105*(47), 18221–18225. doi:10.1073/pnas.0809952105

Franklin, A., Drivonikou, G. V., Bevis, L., Davies, I. R. L., Kay, P., Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *PNAS*, *105*(9), 3221–3225. doi:10.1073/pnas.0712286105

Gallagher, S. (2015). Seeing things in the right way: How social interaction shapes perception. In D. Maxime and Th. Breyer (Eds.), *Normativity in perception* (pp. 117–127). London: Palgrave Macmillan.

Gibson, E. J. (1953). Improvement in perceptual judgments as a function of controlled practice or training. *Psychological Bulletin*, *50*, 401–431.

Gibson, E.J. (1963). Perceptual learning. *Annual Review of Psychology*, *14*, 29–59.

Gibson, E. J. & Rader, N. (1979). Attention – The perceiver as performer. In G. Hale (Ed.), *Attention and cognitive development* (pp. 1–21). New York, NY: Plenum Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: toward an ecological psychology* (pp. 67–82). Hillsdale, NJ: Lawrence Erlbaum.

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

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *PNAS*, *103*(2), 489–494. doi:10.1073/pnas.0509868103

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language*, *105*(2), 91–98. doi:10.1016/j.bandl.2007.06.001

Harris, R. (1980). *The language-makers*. London: Duckworth.

Harris, R. (1990). On redefining linguistics. In H. Davis & T. J. Taylor (Eds.), *Redefining Linguistics*. London/New York, NY: Routledge.

Harris, R. (1998). *Introduction to integrational linguistics*. Pergamon, Oxford.

Harvey, M. I. (2015). Content in languaging: why radical enactivism is incompatible with representational theories of language. *Language Sciences*, *48*, 90–129.

Haugeland, J. (2013). *Dasein disclosed*. Cambridge, MA.: The Harvard University Press.

Hurley, S. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press.

Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism. Basic minds without content*. Cambridge, MA.: The MIT Press.

Hutto, D. D., & Myin, E. (2017). *Evolving enactivism: Basic minds meet content*. Cambridge, MA.: The MIT Press.

Ingold, T. (2001). From the transmission of representation to the education of attention. In: H. Whitehouse (Ed.), *The debated mind: Evolutionary psychology versus ethnography* (pp. 113–153). New York, NY: Berg.

Jackendoff, R. (1996). How language helps us thinks. *Pragmatics & Cognition*, *4*(1), 1-34. doi:10.1075/pc.4.1.03jac

Ježek, E. (2016). *The lexicon: An introduction*. Oxford: Oxford University Press.

Kiverstein, J. & Clark, A. (2009) Introduction: Mind embodied, embedded, enacted: One church or many? *Topoi*, *28*, 1–7. doi:10.1007/s11245-008-9041-4

Kiverstein, J. & Rietveld, E. (2018). Reconceiving representation-hungry cognition: an ecological-enactive proposal. *Adaptive Behavior*, 1–17. doi:10.1177/105971231877277

Kravchenko, A. V. (2007). Essential properties of language, or, why language is not a code. *Language Sciences*, *29*, 650–671. doi:10.1016/j.langsci.2007.01.004

Kravchenko, A. V. (2009). The experiental basis of speech and writing as different cognitive domains. *Pragmatics & Cognition*, *17*(3), 527–548. doi:10.1075/p&c.17.3.03kra

Kukla, R. (2017). Ostension and assertion. In: Z. Adams and J. Browning (Eds.) *Giving a damn – Essays in dialogue with John Haugeland* (pp. 103–130). Cambridge, MA: The MIT Press.

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

Locke, J. (2004). *An essay concerning human understanding*. R. Woolhouse (Ed.). London: Penguin Books. (Original work published in 1690)

Love, N. (1990). The locus of languages in a redefined linguistics. In: H. G. Davis & T. J. Taylor (Eds.), *Redefining Linguistics* (pp. 53–118). London/New York, NY: Routledge.

Love, N. (2004). Cognition and the language myth. *Language Sciences*, *26*, 525–544. doi:10.1016/j.langsci.2004.09.003

Love, N. (2007). Are languages digital codes? *Language Sciences*, *29*, 690–709. doi:10.1016/j.langsci.2007.01.008

Love, N. (2017). On languaging and languages. *Language Sciences*, *61*, 1–35. doi:10.1016/j.langsci.2017.04.001

Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in Psychology*. doi:10.3389/fpsyg.2012.00054

Lupyan, G. (2016). The centrality of language in human cognition. *Language Learning*, *66*(3), 516–553. doi:10.1111/lang.12155

Lupyan G., & Spivey, M.J. (2010). Redundant spoken labels facilitate perception of multiple items. *Attention, Perception, & Psychophysics*, *72*(8), 2236–2253. doi:10.3758/BF03196698

Lupyan, G. & Swingley, D. (2012). Self-directed speech affects visual search performance. *The Quarterly Journal of Experimental Psychology*, *65*(6), 1068–1085. doi:10.1080/17470218.2011.647039

Markman, E.M. (1981). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In: S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought Interrelations in development*, 72–107. Cambridge: Cambridge University Press.

Maturana, H. R. (1978). Biology of language: The epistemology of reality. In: Miller, G. A. & Lenneberg, E. (Eds.), *Psychology and Biology of Language and Thought*. New York, NY: Academic Press.

Matthews, R., (2003). Does linguistic competence require knowledge of language? In: Barber, A. (Ed.), *Epistemology of Language* (pp. 187–213). Oxford University Press, Oxford.

Myin, E. (2016). Perception as something we do. *Journal of Consciousness Studies*, *23*(5-6), 80-104.

Noë, A. (2012). *Varieties of presence*. Cambridge, MA: Harvard University Press.

Ong, W. J. (1982). *Orality and literacy*. London/New York, NY: Routledge.

O'Regan, J. K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 883–917. doi:10.1017/S0140525X01000115

Papafragou, A. (2005). Relations between language and thought: Individuation and the count/mass distinction. In H. Cohen & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (pp. 256–272). Oxford: Elsevier.

Pilling, M., Wiggett, A., Özgen, E., & Davies, I.R.L. (2003). Is color "categorical perception" really perceptual? *Memory & Cognition*, *31*(4), 538–551. doi:10.3758/BF03196095

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York, NY: William Morrow and Company.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. London: Penguin.

Pitchford, N. J. & Mullen, K. T. (2002). Is the acquisition of basic-colour terms in young children constrained? *Perception*, *31*, 1349–1370. doi:10.1068/p3405

Proffitt, D. R., & Linkenauger, S. A. (2013). Perception viewed as a phenotypic expression. In: W. Prinz, M. Beisert, & A. Herwig (Eds.), *Action science: Foundations of an emerging discipline* (pp. 171 – 197). Cambridge, MA : MIT Press.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA.: The MIT Press.

Rączaszek-Leonardi, J. (2009). Symbols as constraints: The structuring role of dynamics and self-organization in natural language. *Pragmatics & Cognition*, *17*(3), 653–676.

Rączaszek-Leonardi, J. (2011). Language as a system of replicable constraints. In H. Pattee (Ed.), *Laws, language and life: Howard Pattee's classic papers on the physics of symbols* (pp. 295–333). Dordrecht: Springer.

Rączaszek-Leonardi, J. (2016) How does a word become a message? An illustration on a developmental time-scale. *New Ideas in Psychology*, *42*, 46–55. doi:10.1016/j.newideapsych.2015.08.001

Rączaszek-Leonardi, J., & Kelso, J. A. S. (2008). Reconciling symbolic and dynamic aspects of language: Toward a dynamic psycholinguistics. *New Ideas in Psychology*, *26*(2), 193–207.

Rączaszek-Leonardi, J., Nomikou, I., & Rohlfing, K. J. (2013). Young children's dialogical actions: the beginnings of purposeful intersubjectivity. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 210–221. doi:10.1109/TAMD.2013.2273258.

Reddy, M. J. (1993). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and Thought (2nd Edition)* (pp. 164–201). Cambridge, MA: Cambridge University Press.

Reed, E. S. (1995). The ecological approach to language development: A radical solution to Chomsky's and Quine's problems. *Language & Communication*, *15*(1), 1–29. doi:10.1016/0271-5309(94)E0010-9

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology (1st ed.)*. New York, NY: Oxford University Press.

Regier, T. & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Science*, *13*(10), 439–446. doi:10.1016/j.tics.2009.07.001

Rietveld, E. & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological Psychology*, *26*(4), 325–352. doi:10.1080/10407413.2014.958035

Roberson, D. & Davidoff, J. (2000). The categorical perception of colors and facial expressions- the effect of verbal interference. *Memory & Cognition*, 28(6), 977–986. doi:10.3758/BF03209345

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). The development of color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General*, *133*(4), 554–571. doi:10.1037/0096-3445.133.4.554

Roberson, D., Davies, I. R. L., & Davidoff, J., (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, *129*(3), 369–398. doi:10.1037//10 096-3445.129.3.369

Rosch Heider, E., & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology*, *3*, 337–354.

Runeson, S. (1977).On the possibility of "smart" perceptual mechanisms. *Scandinavian Journal of Psychology*, *18*, 172–179.

Ryle, G. (1945a). Knowing how and knowing that. *Proceedings of the Aristotelian Society, New Series*, *46*, 1–16. http://www.jstor.org/stable/4544405

Ryle, G. (1945b) Philosophical arguments. In G. Ryle (2009) *Collected Essays 1929-1968*, 203-221. New York, NY: Routledge.

Ryle, G. (2009). *The concept of mind 60th anniversary edition*. J. Tanney (Ed.). London/New York, NY: Routledge. (Original work published in 1949)

Stewart, J. (2010). Foundational issues in enaction as a paradigm for cognitive science: From the origin of life to consciousness and writing. In J. Stewart, O. Gapenne, & E. Di Paolo (Eds.), *Enaction: Towards a New Paradigm for Cognitive Science* (pp. 1–32). Cambridge, MA: The MIT Press.

Suegami, T., Aminihajibashi, S., & Laeng, B. (2014). Another look at category effects on colour perception and lateralisation. *Cognitive Processing*, *15*(2), 217–226. doi:10.1007/s10339-013-0595-8

Tanney, J. (2015) Gilbert Ryle. *The Stanford encyclopedia of philosophy (Spring 2015 Edition)*, E. N. Zalta (Ed.), URL = <https://plato.stanford.edu/archives/spr2015/entries/ryle/>.

Taylor, T. J. (1990). Normativity and linguistic form. In H. G. Davis & T. J. Taylor (Eds.), *Redefining Linguistics* (pp. 118–148). London/New York, NY: Routledge.

Taylor, T. J. (1992). *Mutual misunderstanding: Scepticism and the theorizing of language and interpretation*. Durham, NC/London: Duke University Press/Routledge.

Taylor, T. J. (2000). Language constructing language: the implications of reflexivity for linguistic theory. *Language Sciences*, *22*, 483–499. doi:10.1016/S0388-0001(00)00016-4

Taylor, T. J. (2010). Where does language come from? The role of reflexive enculturation in language development. *Language Sciences*, *32*, 14–27. doi:10.1016/j.langsci.2008.12.014

Taylor, T. J. (2013). Calibrating the child for language: Meredith Williams on a Wittgensteinian approach to language socialization. *Language Sciences*, *40*, 308–320. doi:10.1016/j.langsci.2013.07.002

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thierry, G., Athanasopoulos, P., Wigget, A., Dering, B., & Kuipers, J. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *PNAS*, *106*(11), 4567–4570. doi:10.1073/pnas.0811155106

Turvey, M. T., Shaw, R., Reed, E., & Mace, W. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn. *Cognition*, *9*, 237–304.

Van den Herik, J. C. (2017). Linguistic know-how and the orders of language. *Language Sciences*, *61*, 17–27. doi:10.1016/j.langsci.2016.09.009

Van Dijk, L., Withagen, R., & Bongers, R. M. (2015). Information without content: A Gibsonian reply to enactivists' worries. *Cognition*, *134*, 210–214. doi:10.1016/j.cognition.2014.10.012

Van Elk, M., Slors, M., & Bekkering, H. (2010). Embodied language comprehension requires an enactivist paradigm of cognition. *Frontiers in Psychology*, *1*(234), 1–9. doi:10.3389/fpsyg.2010.00234

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*, 615–628.

Varela, F. J., Thompson, E., Rosch, E., (1991). *The embodied mind: cognitive science and human experience*. Cambridge, MA: The MIT Press.

Verbrugge, R. R. (1985). Language and event perception: steps toward a synthesis. In W. H. Warren, & R. E. Shaw (Eds.), *Persistence and change. Proceedings of the first international conference on event perception* (pp. 157–194). Hillsdale, NJ: Lawrence Earlbaum Associates.

Voloshinov, V.N. (1930/1973). *Marxism and the philosophy of language* (L. Matejka & I. R. Titunik, Trans.). New York, NZ/London: Seminar Press.

Whitehead, A. N. (1929). *Process and reality*. New York, NZ: Harper.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, *4*, 1–13. doi:10.3389/fpsyg.2013.00058

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. doi:10.3758/BF03196322

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *PNAS*, *104*(19), 7780–7785. doi:10.1073/pnas.0701644104

Wittgenstein, L. (1953/2009). *Philosophical investigations (Revised 4th ed.)*. Oxford: Wiley-Blackwell.

Zhu, Q., & Bingham, G. P. (2010). Learning to perceive the affordance for long-distance throwing: smart mechanism or function learning? *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 862–875. doi:10.1037/a0018738

Zukow-Goldring, P. & Ferko, K. R. (1994). An ecological approach to the emergence of the lexicon: Socializing attention. In V. John-Steiner, C. Panofsky, & L. Smith (Eds.), *Sociocultural approaches to language and literacy: an interactionist perspective* (pp. 170–194). Cambridge, MA: Cambridge University Press.