

Elsbeth Jajdelska*

Being there yet not there: why don't embodied responses to literary texts jar with one another?

DOI 10.1515/jls-2016-0002

Abstract: Language and literature can stimulate the embodied resources of perception. I argue that there is a puzzle about why we experience sequences of these embodied responses as integrated and coherent, even though they are not anchored in space and time by a perceiving body. Some successions of embodied representations would even be impossible in real world experience, yet they can still be experienced as coherent and flowing in response to verbal texts. One possibility is that embodied responses to language are fleeting; they need not be integrated because they do not depend on, or relate to, one another as they would in perception. Yet it is the potential for embodied representations to linger and connect with one another which underlies new and persuasive embodied literary theories of vividness, narrative coherence and metaphor comprehension. Another possibility is that readers anchor their embodied representations in a notional human body, one endowed with super-human powers, such as omniscience. But this account relies on implausible, post hoc explanations. A third possibility is that integrating embodied representations produced by language need be no more problematic than integrating the deceptively patchy information harvested from the environment by perception, information which gives rise to an experience of the world in rich and continuous detail. Real world perceptual cues, however, sparse though they might be, are still integrated through grounding in specific points in time and space. To explain the integration of embodied effects, I draw on sensorimotor theories of perception, and on Clark's suggestion (1997, *Being There*. Cambridge, MA: MIT Press) that language can be understood as an additional modality. In this light, the embodied simulations generated by literary texts can be integrated through patterning in a high dimensional, vector space neural architecture, a patterning which recalls real world experience but is specialised to the sustained experience of language itself. This account can help us understand what makes literary experience distinctive and unique.

*Corresponding author: **Elsbeth Jajdelska**, School of Humanities, University of Strathclyde, Glasgow, G4 ONT, UK, E-mail: elsbeth.jajdelska@strath.ac.uk

Keywords: embodied, literary, text, phenomenology, language, perception, sensorimotor, high dimensional, vector space

1 Introduction

Since the discovery that goal oriented motor areas in macaque monkey brains are stimulated by perceiving goal oriented actions (Di Pellegrino et al. 1992), theories of cognition as a form of simulation have gained currency (Rizzolatti and Craighero 2004). With the appearance of evidence for a human mirror neuron system (Buccino et al. 2004), mirror neuron activity has been identified or hypothesised in a range of domains, including the motor actions supporting emotional expression (Dapretto et al. 2006; Nishitani and Hari 2002) and, in aesthetic experience, the motor actions that produce sculpture (Freedberg and Gallese 2007) and music (Overy and Molnar-Szakacs 2009).

Neural simulations in response to hearing and reading language include embodied responses to the motor activity of language production (Fadiga et al. 2002; Morin and Grèzes 2008). But there is also evidence for embodied simulations of verbal content (Kuzmičová discusses the relationship between motor resonance of speech production and content resonance, 2013). Hearing action verbs like “kick”, for example, can stimulate similar brain areas to either performing or seeing the relevant action (Hauk et al. 2008), and hearing concrete nouns like “apple” can stimulate similar brain areas to seeing the relevant object (Reddy et al. 2010; Kosslyn et al. 2005). At the level of discourse, there is evidence that verbal representations of goal-oriented actions can activate brain regions associated with goals (Speer et al. 2009). These brain imaging findings build on earlier and continuing behavioural experiments which indicate that language, including metaphorical language, stimulates content-relevant perceptual simulations (Zwaan and Radvansky 1998; Glenberg and Kaschak 2002; Gibbs 2006; Horton and Rapp 2003).

These findings have challenged traditional modular views of language processing as the manipulation of brittle, amodal symbols (Fodor 1983; Swinney 1979; on the brittleness of classical symbol systems, see Clark, 2016, 24, 162). They also complement recent work on the temporal processing of language, which challenges the modular account of language processing as a linear, feed forward, hierarchical sequence with only minimal feedback mechanisms (Van Berkum et al. 2005; Pickering and Garrod 2013). There has been considerable debate over whether embodied responses to language can or do constitute language comprehension (Barselou 1999; Mahon and Caramazza 2008; Hauk and Tschentscher 2013; Zwaan 2014). But approaches like Pickering and Garrod’s suggest we may

not need to choose between embodied and amodal accounts of language processing (Pulvermüller 2013); that instead of the linear flow-charts of modular accounts, language can engage a suite of simultaneous interactions, recalling Clark's account of predictive processing as a unified computation engaging all the available information (Clark 2016: Ch. 4), or the dynamical mathematical systems underwriting Spivey's *Continuity of Mind* (2007).

In literary studies, evidence for perceptual and motor simulations of verbal content are now being recruited to solve some old literary puzzles. For Mar and Oatley, simulation can explain the evolutionary, social function of fiction, as a method to simplify, abstract and compress real life experience and enable experiential learning and enhanced empathy (Mar and Oatley 2008; Oatley 2011). For Zwaan, embodied simulations can explain narrative immersion (1999). Fine-grained work, often by literary scholars, has considered the ways that texts can differentially engage different cognitive *processes* rather than just choosing between amodal or perceptual symbols. Kukkonen has identified the potential for Bayesian processes of prediction and inference to explain the power of fiction to moderate affect and generate immersion (Kukkonen 2014a, 2014b). Troscianko, Kuzmičová and Jajdelska variously consider the relationships between methods of description, mental representations (above or below conscious awareness) and vividness, with Troscianko suggesting that “cognitive realism”, for example, the extent to which a text engages the machinery of real world cognition, can influence reader's responses to a text (Troscianko 2014b; Jajdelska, Butler, Kelly, McNeill and Overy, 2010,; Kuzmičová 2012). Finnigan shows that descriptions of scenes and rooms are generally structured by the subpersonal “visual primitive” (in Clark's words, a “well-trained feedforward sweep”) used for visual perception of a new environment, rather than by the feature by feature analysis of later, conscious scrutiny (Finnigan 2013a, 2013b; Clark 2016: 51–52). Finnigan also identifies a few writers, such as Virginia Woolf, who seem to work against this cognitive sequence (2013a, 2013b). Jajdelska similarly identifies Nabokov as a writer who can work against as well as with cognitive processes, including limbic system responses to food, drugs, sex and problem solving (Jajdelska 2016).

The idea of literature as simulation, then, need not imply a simplistic “virtual reality” account of verbal art. I wish to build on the work of Troscianko, Kukkonen, Finnigan and Kuzmičová in this respect, moving away somewhat from the fairly close correspondence between cognition of the world and cognition of literary texts found in the work of Zwaan, Mar and Oatley. In what follows, I wish to suggest that language both exploits the embodied resources of cognition and inevitably reconfigures them through very simple and fundamental differences between being *in* the world and reading or hearing *about* the world. These differences, this reconfiguration, mean that even texts

representing banal experience in banal ways can routinely create experiences which are rare or even impossible in real world cognition. The most important distinctions between literary and other kinds of experience, I suggest, go beyond empathy with outgroups (Mar and Oatley 2008, 181), or *Alice in Wonderland* counterfactuals, to let us experience something like “Being There”, to use the title of a pioneering work on embodiment (Clark 1997), but without the physical and social constraints of *actually* being there. Verbal experience can, for example, combine allocentric and egocentric points of view differently from perception (Avraamides et al. 2012). It can let us experience interactions between different modalities which would not be selected for attention in real world experience as relevant to the perceiver’s current needs and goals (Clark 2016: 53–83). It may be time for embodied approaches to literature to start exploring the ways in which literature employs cognitive resources *differently* from real world cognition.

One major potential difference between being there, and reading or hearing about being there, is in the way embodied representations might be combined. Language has the potential to manipulate embodied effects in sequences and at speeds that have no exact counterpart in perceptual experience. The body in space can integrate perceptual information through the speeds and times at which it arrives. But verbal information cannot mimic the timings of touch, sight and sound associated with the act of, say, throwing a stone into a pond. Yet we generally experience the different embodied responses to language as coherent with one another; we do not routinely experience language as an incompatible and puzzling sequence of distinct embodied experiences.

A second major difference is in the current needs and goals of a reader or hearer compared with those of a body embedded in an environment. For the body embedded in an environment, action and attention are governed by the needs to stay physically safe, to engage with social norms, and to pursue short and long term goals. The selection of information directed by a text to a reader can differ radically from the selection of information directed – by the subpersonal processes governing attention, perception and action – to the body in space.

In what follows I start by considering whether embodied representations need to be integrated at all, and suggest that they do. I then identify some general differences between linguistic and perceptual input. I go on to look at the two differences between language and perception identified above – temporal resolution and the physical, social and goal-directed needs of the body in time and space – in more depth using extracts from literary texts. I go on to suggest two possible explanations of how hearers and readers do in fact integrate these embodied simulations, and then identify problems with these accounts. I conclude with an alternative characterisation of language experience. In this account,

language is like an additional modality. Texts supply a flow of cues that can create what would be rare or impossible combinations of embodied experience in real world perception. The reader or hearer identifies statistical regularities in their own responses to this verbal flow. The pattern of regularities we can find in literary experience, while it recalls and overlaps with those of real world perception, is also distinct to the experience of language.

2 Integrating embodied responses to a text

Reading or hearing a text can be, and frequently is, experienced as seamless. As the reader progresses through a ghost story, for example, an account of a room as dark and empty, and a subsequent account of an unexplained noise in that room, may be experienced as an unbroken sequence. This can be manifested physiologically; perhaps the reader's heart rate increases in synchrony with rising goosepimples when she reads about the room, and continues to increase in synchrony with the text about the unexplained noise (Drobes and Tiffany 1997; Zwaan and Truitt 1998; I alternate between gendered pronouns throughout this article). In perception and action, the timings and meanings of mental representations are coordinated and fine tuned by the anchor of a perceiving/acting body in space and time. The reader or hearer of language cannot anchor embodied representations in this way. The physiological *consequences* of integrated embodied representations – the heart rate and the goosepimples – *can* be coordinated with one another through the reader's body. But the reader's body cannot serve as a temporal or spatial anchor for sequences of representations that do not relate to his environment. Why, then, does he not routinely experience a sense of disjunction between the text's effects?

One potential explanation is that comprehension of the text involves embodied simulation of the situation as a whole, and not just of its discrete elements. The psychologists Zwaan and Radvansky have developed an influential account of narrative "situation models", in which texts can be treated as "a set of processing instructions on how to construct a mental representation of the described situation" (Zwaan and Radvansky 1998: 162). In this light, the narrative of the dark, empty room and the sudden noise can be modelled as a single event, of the kind that might give rise to goosebumps and a faster heart rate. So even though the micro embodied experiences – being present in a room, hearing the sudden noise – cannot be anchored and integrated in a perceiving body at the scene, they can form a macro experience which is independent of local details of time and place (Zwaan 1999; Speer et al. 2007).

One potential problem with this argument is circularity; embodied responses in this account both emerge from, and help to build, the situation model. In their 1998 account of situation models, Zwaan and Radvansky initially characterised them as “rather abstract representations”, rather than “movies in the head”, because constructing a (conscious) mental image “requires extensive and effortful processing” which would not be compatible with swift textual comprehension. At this point, their theory was not committed to an explanation of emotion in response to narrative texts. In his 2009 account of immersion, however, Zwaan wondered if physiological responses to narrative texts can be explained by a role for Barsalou’s perceptual symbols and embodied representations (1999) in building situation models.

If it is the situation model that integrates the embodied experiences in the absence of a perceiving body, then the situation model is a prerequisite for the embodied responses. But if the embodied responses are a prerequisite for the situation model, then they would need to be integrated with one another already. We will return to similar arguments in more detail below.

A second potential solution to the integration puzzle draws on a more traditional, amodal and modular view of language comprehension. On this view, while the separate elements in a sentence or discourse – the words, phrases and clauses – *do* need to be integrated by hearers and readers, the embodied representations generated by these elements do not (Swinney 1979; Frazier 1987). The embodied responses, conscious or unconscious, could then be seen as fleeting and contingent effects, epiphenomena independent of comprehension (comparable to mental imagery in Pylyshyn’s account; 1973). This explanation comes at the price, though, of recent progress in explaining literary experience through different kinds of embodied language processing. Zwaan (1999), Troscianko (2014b), Jajdelska (2016) and others assign a central role to the manipulation of unconscious embodied cognitive mechanisms in literary experience. Scarry’s account of literary experience, for example, assumes not only that mental imagery is important, but that sequences of images interact and integrate to create an effect greater than the sum of the parts (Scarry 1999: 6–9, 101–102). Recent work on embodied metaphor processing suggests that even inconspicuous, apparently banal metaphors can rely on the integration of distinct embodied representations (Gibbs 2006). For example, an expression like “grasp the concept” is more readily understood if hearers or readers imagine making a grasping movement. This suggests that the embodied representation of “grasp” is not necessarily a fleeting and isolated effect, but can be, in at least some cases, integrated with other effects to generate meaning, even though there is no real world analogue for grasping a concept.

Having dealt with these initial objections, in what follows I illustrate in more depth the problems associated with treating sequences of embodied responses as overly faithful simulations of real world experience.

3 Some differences between language experience and perception

Language and perception provide information flows with distinct profiles. Sensitivity to the distinctions between experiencing language and experiencing corresponding scenes in the real world can help us understand what makes literature itself a distinctive experience. So while some of the points below may seem, at first glance, too obvious to point out, their implications deserve reflection.

One difference between language and real world perception is speed. It takes much longer to say “the red apple” than to perceive a red apple. And this difference in speed is not constant: “the purple car” takes longer to say than “the red car”, but perception of purple cars is not slower than perception of red cars. A second difference is that linguistic cues are public and directed by an agent. If we are both looking at the same scene, I cannot be certain which cues in the scene are attracting your attention; you might be attentive to the apple or you might be oblivious to it. But I can be sure that we have both heard a voice say “the red apple”, and that both of us have inferred that the voice belonged to an agent. A third difference lies in the relationship between a linguistic cue and its interpretation. This relationship is unstable in different ways from the relationship between perceptual cues and their interpretation. Again, this may seem too obvious to point out to linguists and literary scholars, but languages are subject to systemic changes with internal and external causes which can potentially operate independently of any changes in a mental representation of a percept. At an earlier period in the history of English, for example, “meat” narrowed its meaning from “food” in general, to “flesh” in particular, but there was no corresponding change in the perceptual cues that distinguish flesh from other kinds of food. Fourthly, linguistic information comes in bundles that integrate information from distinct modalities with different weightings than perception does. The word “apple” incorporates, or triggers, information relating to vision, taste, touch and smell, information which can be drawn out by modifiers, as in “the crunchy red apple”. Single modality perceptual cues are also interpreted via higher level representations which are not modal-specific. But the balance of low level cues which feeds in to those high level representations will be contingent on where I am standing relative to the apple; whether I

can smell, taste or touch it and so on. Linguistic information also integrates lower level perceptual cues with higher level interpretations differently. A word like “sunlight” bundles “light”, potentially corresponding to a low level perceptual cue which affects even babies, who have minimal high level mental representations, with “sun”, a high level concept with a place in a complex world model of causes and effects. In real world experience, of course, low level information like “light” is interpreted using high level concepts like “sun”, but in the case of the verbal experience, both “light” and “sun” are processed at the level of language, and therefore both as high level representations (see Reddy et al. 2010 on retinotopic versus ventral temporal cortex in imagery derived from language). The way low level cues and high level representations are bundled in real life cognition, then, cannot be fully simulated in the bundling of embodied responses to language.

Similarly, processing a clause like “the room was dark” cannot consistently be matched up with a specific level of light. But perception responds to precise levels of light immediately. Patterns of attention and depth of processing are also different. Salient *content* in perception might be trumped by salient *sentence position* in language (Sanford and Emmott 2013: 103–131). Finally, language can provide information about absent scenes and stimuli, which perception cannot.

We have good reason to be cautious, then, about making too close a comparison between the experience of real world cognition and the experience of a verbal representation of the world, even if we are happy to commit, as I am, to a central role for embodied resources in responses to literature. It is time to explore in more depth the importance of one important distinction between reading or hearing literature and perceiving relevant content, again a distinction almost too banal to seem worth discussing – the reader’s or hearer’s absence from the scene described.

4 Embodied responses and the reader’s bodily absence

Consider the opening lines of Keats’s “Ode on Melancholy”: “No, no, go not to Lethe, neither twist/Wolf’s-bane, tight-rooted for its poisonous wine;/ Nor suffer thy pale forehead to be kissed/By nightshade, ruby grape of Proserpine” (Keats 1970: 539). Suppose that these lines generate a sequence of perceptual simulations relating to: visiting the river of forgetfulness; twisting the extract from a plant’s root; anointing the forehead with extract of nightshade. It is fairly obvious that real world perceptual counterparts for these simulations could

not be experienced *simultaneously* by one body. And if they were experienced *sequentially* by one body, each experience would affect the next; hands wet from the river would slip on the twisted root, fingers rough from twisting might sting with the extract of nightshade. So while each distinct experience arising from language might be “embodied” in the reader, the integration of the three experiences into a single experience of the poem’s opening lines cannot be anchored in a body in quite the way that real world experience is.

The same is true of texts that develop a scene in a single location and time frame. Here, Vladimir Nabokov describes a family celebration in pre-revolutionary Russia:

I see the steam of the chocolate and the plates of blueberry tarts. I note the small helicopter of a revolving samara that gently descends upon the tablecloth, and, lying across the table, an adolescent girl’s bare arm indolently extended as far as it will go (Nabokov 1998 [1969]: 137).

If the reader fully simulated presence at this scene, he would potentially experience, among other things, the stress associated with violating social norms, which discourage close visual attention to body parts, such as the girl’s arm, in social interactions. Careful attention to the samara and the arm might also generate conflict with the routine real world need to monitor the self in social settings, and to attend to hot drinks and conversation. After all, this is not a scene that permits much slow and silent contemplation among its participants. Empirical evidence, discussed above, suggests that the reader will experience embodied effects that relate to the content of this text. But he cannot integrate these various effects through a faithful simulation of being there.

A participant in a real world scene like this *could* combine self-monitoring and attentiveness to social norms with *some* personal or subpersonal awareness of details like the samara and the arm. These details could be swept up in a glance and then felt, or monitored through covert attention, as the participant carried on with eating, drinking and talking. But that is not, I suggest, the experience which the text, as it stands, is constructed to prompt in a reader. Nabokov’s terms are “I see” and “I note”, rather than, say, “I am aware without looking”. If simulation is in play here, it is simulated overt attention, not simulated covert attention. As with Keats, the range and organisation of perceptual simulations potentially stimulated by the text is distinct from the patterns of salience and attention embodied by presence at the scene. Salience and attention when present are governed by social constraints and physical self-protection, neither of which constrain the reader’s responses. Embodied responses to language, then, can simultaneously require the *disembodiment* of absence if they are not to jar when combined. In Sections 5 and 6 below, I will consider, and query, two potential solutions to this puzzle, and then propose, in Section 7, an alternative solution.

5 Omniscient narrators and other superbeings

One solution is familiar to scholars of literature. We can think of the poet or narrator as the hearer's or reader's proxy in experiencing the simulations, and give this proxy superhuman powers, such as omniscience and invisibility. The difficulty here is that new texts seem to require new post hoc additions to the narrator's or poet's superpowers. For the extract from Nabokov, such a narrator might need invisibility (to stare with freedom at the girl's arm), and the power to slow down time (to watch the samara). For the opening of T.S. Eliot's *The Wasteland* these powers need to be supplemented: "April is the cruellest month, breeding/Lilacs out of the dead land, mixing/Memory and desire, stirring/Dull roots with spring rain" (Eliot 2002 [1922]: lines 1–4). Here, a witness to the stirring roots would need to combine the power to speed up time (since roots stir too slowly to see) with the power to live and perceive underground. This seems as good a point as any to conclude that the omniscient proxy's superpowers, attributed post hoc, exceed a coherent concept of the human body altogether.

6 Real world perception also has to integrate fragmented and incomplete information

Robust findings on attention and perception show that though we experience the world as "a rich, highly detailed and continuous space", what perception actually involves is "a series of momentary snapshots encoded while the eyes are still (fixations) and stitched together across periods of blindness as the eye moves (saccadic eye movements)" (Smith and Henderson 2008; Henderson 2003; Grimes 1996). Perception is "gappy"; without being aware of it, we hypothesise contents for the gaps between fixations by drawing on our experience (in the language of predictive processing theorists, we predict our environments and correct only the errors which are relevant to our current goals; Clark 2016, Ch. 2). The surface of the apple is perceived as red, even though the play of light on its surface produces patches that in other contexts would not seem red (Maloney 1999). In face perception, bottom up cues alone are not enough to let us recognise a familiar face from an unfamiliar angle; recognition is supported by a top down model of the face's underlying structure (Duchaine and Nakayama 2006).

Like perceptual cues, verbal cues supply us with incomplete information about a scene. We do not know if Nabokov is seated on a wooden chair or a stool, or how far he is from his neighbours, for example. So perhaps the method

which integrates fragmentary cues in perception could combine the embodied responses produced by texts? The problem with this solution is that at least some of the integration in real world perception is between sparse, low level, variant information (the varying hues of the apple or the face from varying angles), and high level, invariant information (an invariant model of the causal profiles of apples and faces in relation to light). The variant, low level information may be more sparse than our experience of rich continuous perception leads us to believe. But it is sufficient to ground real world perception in the body's spatial and temporal location. Indeed its power derives precisely from the fact that it is grounded in this way. Without knowing that the cues are coming from a body standing above the face, for example, it would not be possible to relate them to the high level model which explains how light reflects from surfaces at different angles.

The embodied responses evoked by language, on the other hand, draw on higher, invariant processing alone, processing that is independent, in other words, of the body's current location in time and space (Hauk, Johnsrude & Pulvermuller 2004; Reddy et al. 2010). The effects of low level cues from the text itself, the articulation of phonemes, the shapes of the letters, have, in their timing and relation to the body, no effect on the high level embodied representations of content. The method that links high level representations to low level cues in perception, then, cannot alone explain how the disparate disembodied effects generated by texts are integrated.

Perhaps, however, these high level, invariant representations can generate simulations of lower level variant elements, and these could integrate disparate perceptual simulations? The most promising evidence in this direction comes from conscious and unconscious representations of spatial relationships. For example, Kosslyn et al. asked one group of participants to memorise the layout of objects in a picture, and another to memorise a verbal description of the same layout. Both groups showed the same pattern of neural activity, suggesting that texts can stimulate spatial relationships in the same way as perception, and therefore potentially integrate the embodied representations the objects defined by those relationships (2005). Finnigan has shown that a wide range of literary descriptions of scenes are consistent with the spatial organisation of very early visual perception (Finnigan 2013a; Chen 2005). If the top down representations generated by language can be integrated in a spatial framework, it is possible that they do not need to be integrated at the lower levels generated by a body in time and space.

There is a lot to be said for this argument, and there is robust and long standing evidence that representations of spatial *relationships* play an important part in language comprehension (Bransford et al. 1972). So it is possible to build

information from embodied responses to texts into a global spatial model which spans the time period of reading. But this solution suffers from the same problem of circularity as the situation models discussed earlier; if the embodied responses are part of comprehension, then they must be integrated to allow the spatial relationships to be modelled. And if they are not part of comprehension, but merely the side effects of comprehension as modular theorists once argued, then they cannot be recruited to explain affect and immersion in response to literary texts (see above).

7 Patterns of embodied response to the verbal stream

So far I have aimed to establish that embodied responses stimulated by verbal texts must to some extent be combined differently than they are in real world experience. I have considered and rejected some potential explanations for this absence of a jarring response. In this final section, I explore the possibility that the qualities of language itself, the phenomenology of language experience, its “feel” in sensorimotor terms, could integrate embodied representations in response to text with one another.

Clark has suggested that:

Learning a set of tags and labels (which we all do when we learn a language) is, we may thus speculate, rather closely akin to acquiring a new perceptual modality. For like a perceptual modality, it renders certain features of our world concrete and salient, and allows us to target our thoughts (and learning algorithms) on a new domain of basic objects. This new domain compresses what were previously complex and unruly sensory patterns into simple objects (2001: 145).

Thinking about language as an additional modality suggests that embodied reactions to texts could be integrated and made coherent by the distinctive patterns experienced through this distinctive modality. On this reading, language is a systematic disrupter and reorganiser of stored information derived from other modalities, with the power to connect stored experiences, or groups of experiences, that might rarely, if ever, be connected otherwise.

Sensorimotor theories of perception emphasise the patterning of information received through one modality (sight, say) as having a particular “feel” or phenomenology emerging from statistical regularities in interactions between action and perception, in the way that driving a particular car can have a distinctive feel. This could arise, in the case of the car, from the integration of

patterns in the way the gear lever moves, the seat responds to the body, the relationship between pedal pressure and acceleration and so on. The best known articulation of sensorimotor theory was recruited in support of an enactivist account of perception, discussed in relation to literature by Troscianko (O'Regan and Noë 2001; Troscianko 2014a). Troscianko's interest is in O'Regan and Noë's broader theory of perception rather than their account of modality specific phenomenologies. Here, however, I wish to think about language, and specific texts, as having a distinct, embodied phenomenology in a little more depth.

If we think of language as a modality, then the "feel" of that modality, the way that the different states it stimulates are patterned by statistical regularities, would explain how embodied responses to language can be integrated in the absence of a bodily anchor. One way to theorise such patterns is in a vector space architecture, also known as connectionist, or neural, networks (Mordvintsev et al. 2015). In this architecture, different stimuli activate different population codes across cells/neurons by activating or suppressing the cells' activities to different degrees. These population codes capture statistical patterns of similarity and difference between the stimuli across multiple dimensions (Churchland 2012; Rumelhart and McClelland 1986). In this architecture, for example, a given face will activate a pattern of suppression or activation across a very high number of cells, organised in multiple layers, and this pattern will have a family resemblance to the patterns of cells activated not only by other human faces but by the faces of animals and other face-like stimuli (Churchland 2012: 62–77).

The possibility I propose here is that a given verbal text could stimulate sequences of patterns of cell activation which have a family resemblance not only to patterns activated by seeing real world faces, but also to the patterns activated by other *verbal descriptions* of faces, and by verbal descriptions in general. These patterns of similarity and difference, stimulated by variously seeing different kinds of faces, and hearing about different kinds of faces, could be learned and adapted dynamically, in the same way as any other implicit cognitive learning in a vector space (Churchland 2012). Some of these flexible patterns of learning and adaptation in response to language could be specific to a single text, so that we get better at integrating our embodied reactions to that particular text as we proceed through it, just as we get better at integrating our embodied reactions to language in general as we experience more language in infancy.

An example of a set of dimensions on which language might acquire a distinct phenomenology relates to language's status as a form of agent-guided behaviour, governed by, for example, rhetorical goals, or pragmatic cooperative

principles. Other kinds of pattern have been investigated by literary critics, but could apply to non-literary uses of language as well. These include patterns generated from previous experiences of language in social settings, for example as genre, or patterns in linguistic form, for example syntactic parallelism, or rhyme or metre, or patterning in the content of perceptual simulations generated by language, for example in colour or emotion, of the kind identified by Spurgeon's pioneering work on literary imagery (1935). But the patterning that is important in all these cases is between the different states produced in the reader or hearer by the text as stimulus.

In Sections 8 and 9 below, I discuss two potential challenges to this account. The first challenge (in Section 8) comes from empirical findings on narrative processing, and returns to the circularity problem identified earlier in relation to situation models and spatial representations. The second (in Section 9) relates to the experience of text as vivid.

8 Problems with integration through vector space patterning: event structure perception

Work by Speer, Reynolds & Zacks indicates that readers and hearers of verbal narratives pay close attention to changing representations of characters, locations, goals, time and interactions with objects. In this way, they argue, the experience of narrative closely resembles real world “event structure perception” (Speer et al. 2007; Speer et al. 2009; Zacks et al. 2009). These findings are consistent with reports of immersion in narrative; the feeling that the reader or hearer is emotionally engaged in the events, as though she was a participant (Green and Donahue 2009). However, it may be that it is the experience of embodied integration which enables the simulation, and not the reverse, recalling the circularity problem in relation to situation models, discussed above.

To illustrate the problem, here is an excerpt from the Zacks and Speer data, showing how they coded it for their experiment. “1” in a column indicates that the researchers identified a change on the relevant dimension (for example, “time” for a change in time period at this point in the story, “goal” for a change in character's, goals and so on). “0” indicates no change¹:

¹ These materials are available at <http://dx.doi.org/10.1037/a0015305>. They support a number of papers by Zacks and Speer, including: Speer et al. (2007); Speer et al. (2009); Zacks et al. (2009).

		Time	space	object	character	cause	goal
A22	Raymond ran quickly down the terrace.	1	1	0	1	0	0
A23	He stood still	0	0	0	0	1	0
A24	taunting Susan	0	0	0	0	0	0
A25	but not saying anything.	0	0	0	0	0	0

The researchers found neural correlations between identifying a change in the real world in a person's goals, interactions with objects, location and so on, and reading about such changes. This suggests that simulation can explain narrative inference, our ability to fill in information missing from the text, such as Raymond's motive for running down the terrace, or the ways in which he might have taunted Susan (by a defiant expression? By whistling in the air?). But, I suggest, these neural correlates cannot account for the fact that the embodied simulations do not, phenomenologically, jar with one another.

Consider the case of simulating "Raymond ran quickly down the terrace". There are at least two real world analogues here: watching someone run; and running. When watching someone run, a perception of the running as "quick" might draw more heavily on embodied experiences of watching people run at different speeds, measuring the present runner's speed by implicit comparison with runners perceived in the past. When running ourselves, the perception of quickness might draw more heavily on the embodied experience of running, measuring speed by reference to physical effort, adrenalin or discomfort, and to the way the world looks and feels when running at varying speeds. While both sources of experience could be recruited in both kinds of perception, the balance between these two kinds of embodied knowledge will be determined by whether we ourselves are watchers or runners at the moment of perception.

The balance of embodied experiences recruited in response to the verbal description, however, is unlikely to be determined in quite the same way, because in this case the reader or hearer is neither watching nor running. Depending on the ways the text pushes us, our combinations and fluctuations between these two stances are unlikely to have precise real world analogues, a claim with some empirical support from findings on egocentric and allocentric representations of space in episodic memory and narrative processing (Hassabis and Maguire 2007: 303; Avraamides et al. 2012).

Our integration of the embodied representation of "ran quickly" with that of "down the terrace", then, is determined to some extent by *not* being there, even though it draws on embodied real world experience. And perceiving these two representations as an event, or as part of an event, is as much a result of their prior integration as a cause.

9 Problems with integration through vector space patterning: vividness

Readers and hearers of descriptions sometimes report a feeling of “being there”, of perceiving in the real world. But, as Jajdelska et al. (2010), Kuzmičová (2012) and Troscianko (2013) have shown, this kind of vividness is less, not more, likely to arise from descriptions that aspire to a full and faithful recreation of real world counterparts. Here is Nabokov again, as his son collects flotsam on a Riviera beach:

And among the candy-like blobs of sea-licked glass – lemon, cherry, peppermint – and the banded pebbles, and the little fluted shells with lusted insides, sometimes small bits of pottery, still beautiful in glaze and color, turned up (Nabokov 1998 [1969]: 235).

It is possible that the experience of this description as vivid could arise – at least in part – from the *disruption* of sensual childhood memories of the beach rather than their isomorphic reassembly. The description of the glass synaesthetically conflates the experiences of sucking sweets and of collecting sea glass, providing a balance of two inputs to the experience of the glass’s colours, shape and texture that would not normally – perhaps not ever – be combined quite like that in real world perception. Real world perception might give rise to the *thought* that the glass resembles candy, and also to some of the overlapping, potentially synaesthetic, effects. But a body anchored to the beach would not experience these different overlapping representations with the same weightings as the reader, whose embodied resources are exploited while being disembodied with respect to the scene itself. Vividness in this case may be the work of a skilled author pushing together experiences in a way that cancels out some of their content (the glass is not in the mouth, the sweets are not on the beach) so that what is left (taste, texture, colour) is more intense. The fact that verbal descriptions can stimulate vividness, then, can be evidence in favour of a distinct embodied phenomenology of verbal experience, and not against it.

10 Conclusion

The disembodied effects which language can stimulate using embodied resources cannot be integrated, I have argued, through grounding in either a real or a notional body with a location in time and space. Yet they surely are integrated somehow, since readers do not usually find these effects in combination disruptive. I have proposed that if, as Clark suggests, language can usefully

be conceived as a distinct modality, then the embodied experience of language can have a distinct phenomenology, in the same way that sensorimotor accounts assign a distinct phenomenology to distinct modalities. In a high dimensional vector space, patterns of similarity and difference could arise from, for example, linguistic form, synaesthetic overlaps, the author's or speaker's rhetorical goals, speed, tempo and thematic content among many other possibilities. An objection to this account might be the fact that readers sometimes find narrative texts immersive, and descriptive texts vivid; in both cases, the experience is reported to be like that of real world perception. But, I have argued, that experience of vividness strengthens the case for a distinct pattern for the integration of perceptual simulations, while the existence of event structure models in narrative comprehension may depend on it. There is a puzzle about the experience of embodied responses to texts as coherent, and a distinct vector space patterning in response to language is a more plausible solution to that puzzle than a full simulation of bodily anchoring in time and space. Starting from here, I suggest, will give us a richer understanding of how literary texts achieve their effects, and what makes the experience of literature both reminiscent of world experience, and at the same time distinct from it, so that literature can not only represent the world but reinterpret it.

References

- Avraamides, Marios N., Alexia Galati, Francesca Pazzaglia, Chiara Meneghetti & Michel Denis. 2012. Encoding and updating spatial information presented in narratives. *The Quarterly Journal of Experimental Psychology* 66(4). 642–670.
- Barsalou, Laurence W. 1999. Perceptual symbol systems. *Brain and Behavioral Sciences* 22. 577–660.
- Bransford, J.D., J.R. Barclay & J.J. Franks. 1972. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology* 3. 193–209.
- Buccino, G., S. Vogt, A. Ritzl, G.R. Fink, K. Zilles, H.J. Freund, & G. Rizzolatti. 2004. Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron* 42. 323–334.
- Chen, L. 2005. The topological approach to perceptual organisation. *Visual Cognition* 12(4). 553–537.
- Churchland, Paul M. 2012. *Plato's Camera: How the physical brain captures a landscape of abstract universals*. Cambridge, MA: MIT Press.
- Clark, Andy. 1997. *Being there*. Cambridge, MA: MIT Press.
- Clark, Andy. 2001. *Mindware: An introduction to the philosophy of cognitive science*. Oxford: Oxford University Press.
- Clark, Andy. Forthcoming, 2016. *Surfing uncertainty: Prediction, action and the embodied mind*. Oxford: Oxford University Press.

- Dapretto, M., M.S. Davies, J.H. Pfeifer, A.A. Scott, M. Sigman, S.Y. Bookheimer, & M. Iacoboni. 2006. Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience* 9. 28–30.
- Di Pellegrino, G., L. Fadiga, L. Fogassi, V. Gallese & G. Rizzolatti. 1992. Understanding motor events: A neurophysiological study. *Experimental Brain Research* 91. 176–180.
- Drobos, D.J. & S.T. Tiffany. 1997. Induction of smoking urge through imaginal and *in vivo* procedures: Physiological and self-report manifestations. *Journal of Abnormal Psychology* 106. 15–25.
- Duchaine, B.C. & K. Nakayama. 2006. Developmental prosopagnosia: A window to context-specific face processing. *Current Opinion in Neurobiology* 16. 166–173.
- Eliot, T.S. 2002 [1922]. *The Wasteland and other poems*. London: Faber and Faber.
- Fadiga, L., L. Craighero, G. Buccino & G. Rizzolatti. 2002. Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience* 15. 399–402.
- Finnigan, Elizabeth. 2013a. *A Cognitive Approach to Spatial Patterning in Literary Narrative*. PhD thesis: University of Strathclyde.
- Finnigan, Elizabeth. 2013b. Cognitive landscapes: The systematic construction of narrative space. In Rachel Hamilton, Allison Macleod & Jenny Munro (eds.), *Spaces of (Dis)location*, 22–45. Newcastle: Cambridge Scholars Publishing.
- Fodor, Jerry. 1983. *The modularity of mind*. Cambridge, MA: MIT Press.
- Frazier, L. 1987. Sentence processing: A tutorial review. In *Attention and performance XII: The psychology of reading*. M. Coltheart (ed.). Hillsdale, NJ: Erlbaum.
- Freedberg, D. & V. Gallese. 2007. Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Sciences* 11(5). 197–203.
- Gibbs, R. 2006. Metaphor interpretation as embodied simulation. *Mind & Language* 21. 434–458.
- Glenberg, A.M. & M.P. Kaschak. 2002. Grounding language in action. *Psychonomic Bulletin and Review* 9(3). 558–565.
- Green, M. C. & J. K. Donahue. 2009. Simulated worlds: Transportation into narratives. In K. D. Markman, W. M. P. Klein & J. A. Suhr (eds.), *Handbook of imagination and mental simulation*, 241–256. New York and Hove: Psychology Press.
- Grimes, J. 1996. On the failure to detect changes in scenes across saccades. In K. A. Atkins (ed.), *Perception*, vol.5. Oxford: Oxford University Press.
- Hassabis, Demis & Eleanor A. Maguire. 2007. Deconstructing episodic memory with construction. *Trends in Cognitive Sciences* 11(7). 299–306.
- Hauk, O, I. Johnsrude & F. Pulvermüller. 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41(2). 301–307.
- Hauk, O., M. H. Davis, F. Kherif & F. Pulvermüller. 2008. Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging. *European Journal of Neuroscience* 27. 1856–1866.
- Hauk, O. & N. Tschentscher. 2013. The body of evidence: What can neuroscience tell us about embodied semantics? *Frontiers in Psychology* 4(February). 1–14.
- Henderson, J. M. 2003. Human gaze control in real-world scene perception. *Trends in Cognitive Sciences* 7. 498–504.
- Horton, William S. & David N. Rapp. 2003. Out of sight, out of mind: Occlusion and the accessibility of information in narrative comprehension. *Psychonomic Bulletin and Review* 10(1). 104–110.

- Jajdelska, Elspeth, Christopher Butler, Steve Kelly, Allan McNeill and Katie Overy. 2010. 'Crying, moving and keeping it whole: what makes literary description vivid?' *Poetics Today* 31(3). 433–463.
- Jajdelska, Elspeth. Forthcoming, 2016. 'Obnoxious preoccupation with sex organs': The ethics and aesthetics of representing sex. In *Nabokov's morality play: Ethical problems in his fiction*. Michael Rodgers & Susan Sweeney (eds.). Basingstoke: Macmillan.
- Keats, John. 1970. In Miriam Allott (ed.), *Complete Poems*. London: Longman.
- Kosslyn, S., W. L. Thompson, K. E. Sukel & N. M. Alpert. 2005. Two types of image generation: Evidence from PET. *Cognitive, Affective and Behavioral Neuroscience* 5. 41–53.
- Kukkonen, Karin. 2014a. Bayesian narrative: Probability, plot and the shape of the fictional world. *Anglia* 132(4). 720–739.
- Kukkonen, Karin. 2014b. Presence and prediction: The embodied reader's cascades of prediction. *Style* 48(3). 367–387.
- Kuzmičová, Anežka. 2012. Fidelity without mimesis: On mental imagery from visual description. In Gregory Currie, Petr Kot'átko & Martin Pokorný (eds.), *Mimesis: Metaphysics, cognition, pragmatics*, 273–315. London: College Publications.
- Kuzmičová, Anežka. 2013. *Mental Imagery in the Experience of Literary Narrative: views from embodied cognition*. PhD thesis: University of Stockholm.
- Mahon, B. Z. & A. Caramazza. 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology – Paris* 102. 59–70.
- Maloney, L. T. 1999. Physics-based approaches to modeling surface color perception. In K. R. Gegenfurtner & L. T. Sharpe (eds.), *Color Vision: From genes to perception*, 387–417. Cambridge: Cambridge University Press.
- Mar, Raymond & Keith Oatley. 2008. The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science* 3(3). 173–192.
- Mordvintsev, Alexander, Christopher Olah & Mike Tyka. 2015. Inceptionism: going deeper into neural networks. *Google Research Blog*, 17 June 2015: <http://googleresearch.blogspot.co.uk/2015/06/inceptionism-going-deeper-into-neural.html>, accessed 17 September 2015.
- Morin, O. & J. Grèzes. 2008. What is "mirror" in the premotor cortex? A review. *Clinical Neurophysiology* 38. 189–195.
- Nabokov, Vladimir. 1998 [1969]. *Speak memory*. Harmondsworth: Penguin.
- Nishitani, N. & R. Hari. 2002. Viewing lip forms: Cortical dynamics. *Neuron* 36. 1211–1220.
- Oatley, Keith. 2011. *Such stuff as dreams: The psychology of fiction*. Chichester: John Wiley & Sons.
- O'Regan, J. Kevin & Alva Noë. 2001. What is it like to see? A sensorimotor theory of perceptual experience. *Synthese* 129. 79–103.
- Overy, Katie & Istvan Molnar-Szakacs. 2009. Being together in time: Musical experience and the mirror neuron system. *Music Perception: An Interdisciplinary Journal* 26(5). 489–504.
- Pickering, Martin & Simon Garrod. 2013. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences* 36. 329–392.
- Pulvermüller, F. 2013. How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences* 17(9). 458–470.
- Pylyshyn, Z. W. 1973. What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin* 80. 1–25.
- Reddy, L., N. Tsuchiya & T. Serre. 2010. Reading the mind's eye: Decoding category information during mental imagery. *Neuroimage* 50. 818–825.

- Rizzolatti, Giacomo & Laila Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27. 169–192.
- Rumelhart, D. E. & J. L. McClelland. (eds.). 1986. *Parallel distributed processing*, 2 vols, Cambridge, MA: MIT Press.
- Sanford, Anthony & Catherine Emmott. 2013. *Mind, brain and narrative*. Cambridge: Cambridge University Press.
- Scarry, Elaine. 1999. *Dreaming by the book*. New York: Farrar, Straus and Giroux.
- Smith, Tim & John M. Henderson. 2008. Edit blindness: The relationship between attention and global change blindness in dynamic scenes. *Journal of Eye Movement Research* 2(2). 6, 1–17.
- Speer, Nicole K., Jeffrey M. Zacks & J. R. Reynolds. 2007. Human brain activity time-locked to narrative event boundaries. *Psychological Science* 18. 449–455.
- Speer, Nicole K., J. R. Reynolds, K. M. Swallow & Jeffrey M. Zacks. 2009. Reading stories activates neural representations of visual and motor experiences. *Psychological Science* 20. 989–999.
- Spivey, Michael. 2007. *The continuity of mind*. Oxford: Oxford University Press.
- Spurgeon, Caroline. 1935. *Shakespeare's imagery and what it tells us*. Cambridge: Cambridge University Press.
- Swinney, David A. 1979. Lexical access during sentence comprehension: (re)Consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 18. 645–659.
- Troscianko, Emily T. 2013. Reading imaginatively: The imagination in cognitive science and cognitive literary studies. *Journal of Literary Semantics* 42(2 September). 181–198.
- Troscianko, Emily T. 2014a. Reading Kafka enactively. In Terence Cave, Karin Kukkonen, & Olivia Smith (eds.), *Reading literature cognitively*. Paragraph special issue, 37(1). 15–31.
- Troscianko, Emily T. 2014b. *Kafka's cognitive realism*. Abingdon and New York: Routledge.
- Van Berkum, J. J. A., C. M. Brown, P. Zwitserlood, V. Kooijman & P. Hagoort. 2005. Anticipating upcoming words in discourse: Evidence from ERP's and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31. 443–467.
- Zacks, Jeffrey M., Nicole K. Speer & J. R. Reynolds. 2009. Segmentation in reading and film comprehension. *Journal of Experimental Psychology* 138. 307–327.
- Zwaan, Rolf A. & G. A. Radvansky. 1998. Situation models in language comprehension and memory. *Psychological Bulletin* 123. 162–185.
- Zwaan, Rolf A. & T. P. Truitt. 1998. Smoking urges affect language processing. *Experimental and Clinical Psychopharmacology* 6. 325–330.
- Zwaan, Rolf. 1999. Embodied cognition, perceptual symbols, and situation models. *Discourse Processes* 28. 81–88.
- Zwaan, Rolf. 2014. Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences* 18(5). 229–234.