Draft of piece published in The Mind, the Body and the World: Psychology After Cognitivism?, B. Wallace. A. Ross, J. Booth Davies and T. Anderson (eds.), Imprint Academic. 2007, pp.298-305.

Continuity in Question
An afterword to *Is Language the Ultimate Artefact?*

Michael Wheeler
Department of Philosophy
University of Stirling
FK9 4LA
UK
m.w.wheeler@stir.ac.uk

Is Language the Ultimate Artefact? (henceforth ILUA) was originally published alongside a paper by Andy Clark called *Is Language Special? Some remarks on control, coding, and co-ordination* (Clark 2004). One concern (among others) of the latter paper was to resist the argument of the former. In this short afterword, I shall attempt a counter-response to Clark's resistance. In so doing I hope to reveal, in a new and perhaps clearer way, what the most important issues really are in this (still unresolved) debate.

Let's begin by recalling the pivotal disagreement at the heart of the matter, as identified in the original exchange. Clark's position-defining claim is that language is "an external resource that complements but does not profoundly alter the brain's own basic modes of representation and computation" (Clark 1997, p.198). The issue concerns the representational rather than the computational half of this continuity. The disputed point, reiterated by Clark in his response to me, is that "[the] brain represents [linguistic] structures, of course. But it does so in the same way it represents anything else. They do not re-organize neural routines in any way that is deeper or more profound than might occur, say, when we first learn to swim, or to play volleyball." (Clark 2004, p.720) Now I agree that this claim holds in the case of on-line language use, that is, in the case of language-involving behaviour in which the relevant material symbols (such as printed text or ambient linguistic sounds) are present in the currently accessible environment and may thus form proper parts of a real-time distributed cognitive process.[1] However, I argue that it fails in the case of off-line language use such as linguistic inner rehearsal, in which the relevant material symbols are, by hypothesis, not present in the currently accessible environment. Here (again) is why.

Given a distributed approach to cognition (which is ground shared by Clark and me), how one understands the transition between on-line and off-line language use must be shaped by the general (i.e., non-language-specific) point that, in paradigmatic cases of distributed cognition, adaptive success ensues because, during the actual run-time of the behaviour, certain internal elements become directly causally locked onto the contributing external elements. When our reasoning is off-line, there are, by hypothesis,

[1] The term 'material symbols' is due to Clark.

no such environmental factors onto which the mechanisms concerned could be locked. Nevertheless, the claim that fundamentally the same kind of computational processes are in play might survive, just so long as there exist certain inner surrogates for those missing environmental factors, surrogates which recapitulate certain structural properties of those factors, viz. the ones, whatever they may be, to which the mechanisms concerned are designed so as to be mechanically keyed. So, in the linguistic case, what we need are inner surrogates that realize the critical structural properties of the very linguistic material symbols that support the corresponding on-line language-involving behaviour. Since the critical structural properties will be linguistic ones, the inner surrogates must themselves be linguistically structured (or so I suggested in ILUA). So Clark's claim that linguistic structures "do not re-organize neural routines in any way that is deeper or more profound than might occur, say, when we first learn to swim, or to play volleyball" is violated. A sub-set of language-related cognition requires a transformation in the brain's own basic mode of representation, from one that is essentially non-linguistic in form to one that is essentially linguistic. Here's how Clark describes this alleged predicament. ('My' and 'I' in the following quotation refer to Clark.):

> [On the one hand, my position] means rejecting the idea that language processing requires some very special kind of internal processing and representation. On the other hand, I want to unpack offline cogitation, quite generally, in terms of internal recapitulations of the relevant-but-missing environmental structures. Since the environmental structures, in the linguistic case, are quite patently (perhaps tautologously) structured in a linguaform way, did not I just lose the farm, at least as far as the internal representations are concerned? Perhaps, Wheeler concedes, the *processing* can be unaffected… But the *internal representations* really have been radically re-structured by the need for the offline use of linguistic resources. (Clark, 2004, p.721)

Clark attempts to escape from this potential quandary by suggesting that my argument actually falls short of its intended mark.

> All [Wheeler's] argument shows is that the inner surrogates must amount to a representation, useable off-line, of the relevant environmental structures. But a representation of structure is not thereby… a structured representation. Just as I can represent greenness without deploying a green inner vehicle, so too I can represent a sentence as involving three component ideas (John, loving, and Mary, to stick with the tired old example) without thereby deploying an inner vehicle that itself comprises three distinct symbols exhibiting that articulation. (Clark 2004, p.722)

Now, Clark is surely right to point out that a representation of some X-with-a-certain-structure is not *thereby* an X-structured representation. So he is indeed correct that one can represent a structured sentence without *thereby* deploying an inner vehicle that itself comprises *all the very same elements and structure as that public sentence*. My original argument moved too quickly for its own good. But establishing that some inner surrogate

need not recapitulate *all* the very same elements and structure as are realized by its representational target does not establish that that surrogate need not recapitulate *certain* structural properties of that target. And although it may well be obvious that representing greenness does not necessitate the presence of a green inner vehicle, it is rather less obvious that representing a *missing* structured syntactic object, in a way that secures competence in the relevant cognitive domain, does not necessitate some degree of significant structural recapitulation on the part of the inner element, and that's all my argument needs. So the unresolved question is *how much* structure is needed, and what the *character* of that structure is, in particular cases.

What Clark owes us, I believe, is a developed account of how linguistic inner rehearsal may take place *without* the kind of significant recapitulation of linguistic structure that, I suggest, is lurking in the explanatory wings. Fortunately, a sketch of how such an account might go is on offer. Clark calls it the *cognitive self-stimulation model* of off-line language use (Clark unpublished). So let's see how it fares. As I understand it, the core of Clark's view is that, in off-line language-use, human beings do their thinking using inner *images* of words. The notion of an 'image' should be understood in a wide sense here, so as to include structures with an auditory or multi-modal character, as well as those of a purely visual kind. Thus, according to Clark, in on-line language use we access certain environmental inputs (e.g. the word on the page, the sentence in the air) that stimulate the brain, via sensation, so as to perturb it into different regions of its state space. This account of on-line language use is designed to cohere with the general distributed cognition approach to on-line intelligence. Subsequently, in off-line language use, we 'simply' self-create surrogates for the now-missing inputs, sometimes in the form of potentially observable structures such as audible vocalizations, but often in the form of inner images of the sort just mentioned. Given that these self-created surrogates are designed to preserve only the relevant sensory properties (the shapes that one sees, the sounds that one hears), they do not realize linguistic structure. Moreover, it is unmysterious how they might invoke essentially the same inner processes as are invoked by the environmental inputs in the on-line case. Thus continuity of representational structure and of computational process is preserved.

So why should there be cause to worry about this undeniably attractive story? The first thing to note is that Clark's flagship example of off-line mathematical reasoning (which, as I argue in ILUA, is analogous to the linguistic case) tends to skew one's receptivity to the self-stimulation model. Clark focuses on the example of using imagined Venn diagrams in our heads, and notes that "there is no reason to suppose that… [this requires]… the installation of a different kind of computational device", different, that is, to the one active during cases of on-line reasoning involving Venn diagrams on the page (Clark, 1997, p.199). A cognitive self-stimulation account would seem to have some cogency here. One simply self-creates images of the missing diagrams and deploys the same processing strategies. However, because of the fundamentally spatial, and therefore essentially visual, nature of the reasoning, the Venn diagram example rewards the idea of self-created pseudo-sensory inputs, but is potentially misleading as to the general prospects for the cognitive self-stimulation model, and thus for the kind of continuity that Clark advocates. To see this, consider another, and arguably more central, example of off-line mathematical reasoning, namely performing in one's head the kind of calculation that, in the on-line case, might standardly be tackled using pen and paper and the

machinery of long multiplication, but which under pressure we can perform off-line. Do we really imagine the carryings of numbers that figure as on-the-page manipulations in the on-line case? That's the sort of thing that would seem to be required for cognitive self-stimulation to get any purchase. Phenomenological intuitions may vary, of course, but I'm willing to bet that, for most people, imagined carryings are just not part of our experience here. If this is right, then there is some evidence already to suggest that the door is far from closed on the sort of transformation in representational structure that I have argued is present in the transition from on-line cases of language-involving cognition to their off-line cousins.

So let's now open that door as wide as we can, by reflecting on Clark's own chosen example of an extant cognitive-scientific approach that demonstrates his key point, viz. that a representation of some X-with-a-certain-structure is not thereby an X-structured representation. That example is Elman's dynamical connectionist modelling of language (Elman 1995). The relevant studies here feature what has come to be known as an Elman net, a simple recurrent connectionist network in which the widely used three-layer architecture of input, hidden, and output units is extended to include a group of context units. These units store the activation-values of the hidden units at any one time-step, and then feed that information back to the hidden units at the following time-step. The present state of such a network is thus a function of both the present input and the network's previous state, which, as Elman shows, allows this sort of system to encode sequential information, and thus to succeed at certain prediction-tasks. For example, given a corpus of simple sentences, constructed from a small set of nouns and verbs, and presented so that the only information explicitly available to the network was distributional information concerning statistical regularities of occurrences in the input strings, Elman was able to train a simple recurrent network to predict the cohort of potential word successors under various conditions of use. (A cohort is made up of all the words consistent with a given span of input.)

Subsequent statistical analysis of the network demonstrated that it had achieved its predictive capabilities by inducing several categories of words which were implicit in the distributional regularities in the input data. These induced categories had an implicitly hierarchical structure. Nouns were split into animates and inanimates, sub-categories which themselves were subdivided (into, for example, classes such as humans, nonhumans, breakables, and edibles). Verbs were grouped in the following categories: (a) requiring a direct object, (b) optionally taking a direct object, and (c) being intransitive. Conceptually similar words drive the network into regions of activation space that are close together, so conceptual similarity is captured via position in activation space. The most general linguistic categories, such as 'noun' and 'verb', correspond to large areas of this space, whilst more specific categories and (ultimately) individual words correspond to progressively smaller sub-regions of larger areas. Thus the space implicitly realizes the hierarchical structure described above.

According to Clark, this Elman net is a language-navigating system that represents linguistic structure while failing to realize linguistically structured representations. However, as far as I can see, this judgment can't be right. In my view the statistically visible states realized by this network *do* qualify as inner elements with linguistic structure. Of course, given the nature of the emergent groupings in question (see above), the states at issue do not perhaps obey the traditional syntax-semantics

distinction, but that, in and of itself, doesn't stop those states from being linguistic in character. Its activation space is structured in terms of nouns, verbs, transitivity, intransitivity, and so on. How much more linguistic does it get?

At this point someone might be tempted to point out that, as with any connectionist network of this kind, there's a representational-computational level of description at which we don't see linguistically structured states, but rather weight matrices and patterns of activation values. However, that observation alone can't carry the day. As should be clear from my description of Elman's study, the statistically visible states here do genuine explanatory work Under these circumstances I think we should unhesitatingly proceed to reify those states. And if that's correct, then what we have here is a system that realizes linguistic structure, and *not* a language-navigating system that represents linguistic structure while failing to realize linguistically structured representations.

One further objection that Clark might be tempted to make here is suggested by something else that he says in *Is Language Special?*. He notes there that my position depends on the assumption that it is theoretically possible to separate the structure of the representation from the nature of the associated processing, in such a way that it would be possible to speak of a fundamental transformation in the mode of representation but not in the associated mode of computation. I agree that the argument of ILUA depends on this assumption. Indeed, I'm up front about it (see ILUA footnote 18). But, according to Clark, this assumption is ultimately misguided, since "for a representation to genuinely be structured in a certain way, just IS for the system to be able to operate upon it in certain ways" (Clark, 2004, p.722, footnote 3). But now if representation and computation are intimately co-defined, one might refuse to reify the higher-order states of the Elman net on the grounds that since all the computational processing plausibly goes on at the level of the connection weights and unit activations, the higher-order states in question cannot be the objects of computational processes, and so cannot be representational in character, and so cannot figure in the fundamental representational-computational story. Of course, I'm inclined to reject the claim of intimate co-definition, but rather than argue for that here, I'd like to draw out a consequence of using that claim to resist the reification move. If the proposed alternative condition for reification (being the object of a computational process) were to be applied generally, and if we assume for the moment that connectionism provides a good model of the fundamental character of mind (as Clark does), then the only level at which one could speak of representations at all is at the level of weights and activation values. The present proposal would thereby mandate the elimination from our cognitive ontology of all sorts of higher-order psychological structures that are not reflected directly in the lower-level processing story. Writing blank cheques to eliminativism is not something that I think Clark would want to encourage. So this objection fails too.

To repeat by way of conclusion: The test-case Elman net is *not* a language-navigating system that represents linguistic structure while failing to realize linguistically structured representations. Rather, it is a system that navigates language by itself realizing linguistic structure. Thus it does not provide the kind of evidence that Clark needs in order to resist the argument of ILUA. In my view that argument is still well worth the pages on which it's written.

References

Clark, A. (1997), Being There: Putting Brain, Body, And World Together Again (Cambridge, MA. and London: MIT Press/ Bradford Books).

Clark, A. (2004), 'Is language special? Some remarks on control, coding, and co-ordination, Language Sciences, 26:6, special issue on Distributed Cognition and Integrational Linguistics, ed. D. Spurrett, pp. 717-726.

Clark, A. (unpublished), 'The ins and outs of language'. Paper give at the University of Dundee, February 2006, in dialogue with Michael Wheeler.

Elman, J.L. (1995), 'Language as a dynamical system', in Mind as Motion: Explorations in the Dynamics of Cognition, eds. R. Port and T. van Gelder (Cambridge MA.: MIT Press/ Bradford Books), pp. 195-225.