Language in Iterating Activity: Microcognition Re-membered

Grant Blair and
Stephen Cowley

1.0 Introduction

Radical as its consequences may be, *prima facie* there is nothing implausible about a distributed view of language. Linguistic theory notwithstanding, even casual observation suggests that, during iterated human activities, what we ordinarily call 'language' spreads across brains, bodies and world. Seeking ways of understanding this spread, we urge fresh thinking about verbal labels. Specifically, we suggest that this central part of language is grounded in activity that is social *and* microcognitive. Dynamic and dialogical activity, words spoken, promote remembering. As with many computational resources—music, mathematics, navigation—to exploit labels for remembering is to take part in social as well as sub-personal activity.

Our strategy is incompatible with defining language (or languages) round word-based units or 'linguaform'. Rather than stress internal processes, we highlight that much of the *sense* of utterance-activity arises at and across the boundaries of skin and skull. Much of what we mean and do is, literally, real-time responding to prosodic and visible expression. Clarifying, we use Kirsh and Maglio's (1992; 1994) work on how success (or failure) in computer games exploits real-time activity. Using parallels between a player's microcognitive moves and events of utterance-activity, we both endorse and critique aspects of Clark's (1997; 1998) externalist view of language. Treating utterance-activity as intrinsic to cultural processes that augment the brain's powers, we stress that—during talk—language is *also* embodied activity. This thesis, we claim, throws conceptual light on Cowley's (1994; 1997; 1998; 2001b) empirical work on how prosody functions. Finally, we claim that

microcognitive abilities enable word-based labels to promote remembering. What human bodies and voices do *together* affects how we conceptualise our encultured worlds.

2.0 Cognitive Internalism and Language Models

Influenced by the work of, especially Descartes and Hume, many thinkers take for granted that cognitive processes occur exclusively *within* the brain. Certainly, linguists have generally accepted some version of 'cognitive internalism' without demur. Combining this position with a wish to describe language scientifically, their models often present linguistic *abstracta*—static entities—as units susceptible to internal processing.[1] These views are shared by theorists with perspectives as diverse as those of, for example, Chomsky and Halliday. Such 'process' models can be justified within the standard parameters of the discipline. Not only do they fit the Saussurian ideal of synchronic description but they are consistent with the practice of separating language from so-called 'use'.

For those interested in mind, defining languages as *abstracta* makes linguistic theory compatible with input-output views of what we do with words. Generally, then, what Churchland (1995) first called 'linguaform' has become the main focus in linguistic theory. Remarkably, this is as true of theories positing psychological and biological realities as those which investigate language in relation to a range of cultural, social and even applied issues. Of course, defining language as linguaform has the (dubious) advantage of making language 'processes' independent of other cognitive functions. Thus while empiricists treat language as independent of the brain, rationalists can see it as having a mere 'interface' with the central nervous system. For such theorists sentences and other abstract entities are input and output forms exploiting some kind of language module. Somewhere in the head, it seems, there is a locus where thought is decked out as language and dedicated processes send and receive linguaform mediate information.

[1] Notable twentieth century linguists who have emphasised the external complexity of language include Bloomfield (1933), Harris (1951) and Harris (1981).

Those adopting a process model generally aim to clarify what we know when we know a language. Instead of asking how brains function during, for example, gossiping, worshipping, swearing, doing philosophy, making friends, talking to dogs etc., it is supposed that—during these activities—we process linguaform patterns. Cognitive internalists seek to explain language (or languages) by models of units constituted as sets, systems and formal notations. Dividing syntax from meaning—and/or function from use—they treat *abstracta* as rule governed or, perhaps, as subject to probabilistic patterns. Such descriptions, moreover, are 'tested' on a serial, digital computing device. On process views, then, the linguistic mind is reduced to a putative entity that stores and processes information, generates sentences, and/or construes *a priori* meanings. Even if neurally distributed, language processes are treated as occurring only inside the head. Unlike when brains serve how individuals act in encultured environments, language depends on manipulating *abstracta*. Exploiting our alleged capacities for producing and processing such *abstracta*, they are thought to give rise to output and input that serve brain-internal goals.

2.1 Towards a Distributed View

In the last decade cognitive internalism has been challenged on philosophical, scientific and practical grounds (e.g. Hutchins 1995; Clark 1997; 2001; Rowlands 1999; Hurley 1998; Spurrett this volume; Cowley & Spurrett in preparation). At a theoretical level, as Hutchins shows, working a navigation system depends just as much on other people and external devices as processes within the head. As he has argued, we must acknowledge the importance of 'culturally distributed cognition'. Parallel to this, many have challenged the view that most (or all) inner cognitive processes are appropriately modelled by manipulating internal symbols (e.g. Brooks 1991a; Clark 1997; van Gelder 1997). Reacting against the 'computational theory of mind', many now doubt whether brains use symbols at all. Indeed, even Fodor (2000) emphasises that much of cognition is irreducible to anything that can be described by such models.

As the brain ceases to be seen as an exclusive locus of symbol manipulation, Clark (2001) is proposing that we develop a new vision of human rationality. Set alongside the above-mentioned 'symbol flight' this

demands radical thinking about how, in various time-scales, language is grounded. Rather than define it around *abstracta*, we regard language as a non-determinate, dynamic entity, subject to various kinds of constraints (see Harris 1996). Further, we hypothesise that its emergence depends on a human capacity to exploit particular classes of what Dennett (1991b) calls 'real-patterns'.[2] On this view, what needs to be clarified, then, is how iterated social activity allows us to pick up on these patterns and, in so doing, encourage their spread between world, artifacts and brains. In these broad terms, moreover, we find commonality across the work of scholars whose interests are as different as those Deacon (1997), Dennett (1991a), Hutchins (1995), Churchland (1995) and Clark (1997). All share a concern with how language-behaviour links brains, bodies and world.

In developing a distributed view of language, we stress that that our main cognitive resource is—not linguaform—but a diverse, encultured world. Bodies, unlike serial, digital symbol-processors, adapt to their surroundings and, on occasion, allow humans to adjust the world to their requirements. It is because of this capacity for adjustment that many believe intelligence can only be understood 'from the-bottom up'. Accordingly, theorists are investigating how devices can adapt to their environments and, above all, what microcognitive resources are required. While not seeming 'cognitive', what insects and robots do is today, from a behavioural and computational view, well understood (see Clark 1997). In grounding language, then, we argue that similar microcognitive resources give access to an encultured world. Thus, we examine language—in its simpler manifestations—with respect not just to linguaform (e.g. written signs) but also to constraints that shape 'stand-ins' and their functions.

Dennett (1991) provides a neat analogy in discussion of colour vision. When we see, say, a 'red' object, the colour is neither in the world nor in the head. Rather, a capacity to do what we call 'seeing colour' uses a stand-in that is, ultimately, the outcome of a co-evolutionary arms race between plants and insects. In ways that remain unclear, cultural co-evolution may have led to selection of individuals who *hear* human vocalizations as word-forms. Following von Uexkull (1934) they belong neither to a subject nor the world

[2] Patterns are a regularity in physical phenomena which are by definition compressible and recognisable.

but, rather to the organism's surroundings or *Umwelt*. Further, in some way, these entities are part of its *Innenwelt* (they are stand-ins). Pursuing this view, we play down internal cognition and perception and, instead, explore microcognitive aspects of *utterance-activity*. Highlighting what people do, we stress how labels remind us of things. Singly and jointly they allow us to exploit what previously happened against current events. They allow us to *re-member* experience. By bringing back to mind whatever-there-is (for our community and ourselves), use of labels points towards thinking of language in terms of cross-overs between what people do. This 'in and out view' of talk plays down linguaform while, alluding to Wittgenstein's work, giving a new perspective on varying and iterating activity.

3.0 Language and Labels

Where language is seen in terms of static linguaform patterns, its use is pictured around operations on classes of labelled entities. This is because, if language has *a priori* existence, its manifestations must reiterate forms and functions that—in principle—can be labelled. As noted, the alternative is to regard language as an indeterminate, dynamic system subject to various constraints. Rather than its symbolic nature, what comes centre stage is how language augments a capacity to contextualize experience (see Harris 1997; Cowley, 2001b). In turn, this depends on an inherent reflexivity (see Harris, 1995) that requires us to think, not in terms of 'use', but how the uniqueness of language contributes to individual achievements. Language-mediated events, we think, depend less on labelling (or shared knowledge) than how we use macrosocial, circumstantial and biomechanical constraints (see Harris 1995). Over time, these set off development that can be fitted to a person's (perceived) interests by turning language on itself and the world. With Wittgenstein (1958), we think it is mistaken to treat the capacity to exploit labels as a necessary condition for language. Rather, we see this as deriving from how talk contributes to human practices. During these events persons, things and language get enmeshed in social activities that allow use to begin to exploit labels. Just as red objects help us develop a category 'red', iterating activity (including language) enables us to label linguistic, mental and practical events.

Our view of labelling contrasts with the one that Wittgenstein[3] represents in a long citation from the work of Augustine:

> When they my elders named some object, and accordingly moved towards something, I saw this and I grasped that the thing was called by the sound they uttered when they meant to point it out (PI 1).

For the Saint, to grasp a thing in language depends on how a knowing 'I' connects with a name. What cognitive scientists have called the 'central executive' unites an object seen, a sound heard, and an intention identified. For the label to stick, moreover, the connection must come to be stored by a neural system. In taking a parallel view, Augustine is at pains to see off an obvious objection. Thus, he seeks to clarify how intentions are identified:

> Their intention was shewn by their bodily movements, as it were the natural language of all peoples: the expression of the face, the play of the eyes, the movement of other parts of the body, and the tone of voice which expresses our state of mind in seeking, having, avoiding or rejecting something (PI 1).

If Wittgenstein's demolition of the Augustinian picture is conceptual, we wish to stress that there is also something wrong with its empirical framing. While humans share propensities for expression, individuals and cultures exploit bodies differently. There is no 'natural language of all peoples'. Opposing the implied cognitive internalism, we need an alternative view of how words can be grounded in interpersonal events.

Although Augustine's quotation is the keystone to the *Philosophical Investigations*, the spell of Saint's view remains. Indeed, unless the reflexivity of language is recognised, we think this symbolic picture will continue to mask understanding of language, culture and brain. Until then, theorists will abstract word-forms from what people do (and the sense of their talk) and, treat understanding as an inner process. If they do not appeal to the mysteries of generative grammar, they may persuade themselves that linguistic signs arise

[3] References to *Philosophical Investigations* (1958) are to section numbers in the form (PI section no.)

from parallel choice-making in systems of labelled kinds.[4]

Rejection of the Augustinian model provides a basis for thinking in dynamic terms. Of course, in making this move, we must be wary of rejecting labels. As Wittgenstein puts it, 'It will often prove useful to say to ourselves: naming something is like attaching a label to a thing' (PI 15). What has to be established, then, is a view of labelling having little to do with symbols or *a priori* categories. Rather that prioritise linguaform, we ask how, on a given occasion, a goal-driven agent can act by attaching signs to objects, events, or abstract categories. However, we stress that nothing *binds* subsequent action (or practice) to what we treat as this linguaform. A similar thought is expanded in Wittgenstein's discussion of Augustine's view. In the parable of the builders, we are asked to imagine a simple 'language'. In using this, however, the builders engage in social practices inseparable from gesticulating and pointing. For them, a sign 'means' if and only if other parties know what to *do* with it. It is only insofar as actions are performed against a background of iterated practices that this labelling makes sense. Thus not only must a practice of ostensive teaching (as opposed to definition) already be in place but it would be circular to invoke labels in explaining language *origins*. To grasp how language is grounded, we must trace how words come to replace natural forms of expression. Drawing on Deacon's (1997) work, we may differ from other species in that our iconic and indexical communication eventually also becomes symbolic.[5] As Spurrett and Cowley (2003) show, the natural history of human infants is shaped by how, in a physical and cultural setting, actions exploit caregivers. Another way of thinking about a dynamic approach is found in a story by Marquez (1972). In the town of Macondo, we read, an illness

[4] Halliday and Mattiesson (1999) take just such a view as the cognitive grounding for systemic-functional theory. From our perspective this is no improvement on the static process views associated with what Matthews (1993) terms the American descriptivist tradition: in brief, this is because it still relies on a synchronic system to explain a bogus distinction between language and use.

[5] However, taking the symbolic nature of language for granted, Deacon (1997) posits the existence of a cognitive bottleneck that prevents other primates from developing the relevant abstracta. In response, Cowley (2002) argues that this bottleneck may not be 'in the brain'.

leads to insomnia and a loss of memory. However to combat loss of memory, the townspeople begin to spend night hours labelling things. So they put up a big sign at the entrance to the village saying Macondo and, right next to it, being Catholic, a sign 'God Exists'. This all goes well. Entire households are over-run by little slips of paper, until the people begin to forget what each item is used for. After a time, this leads to the cow being labelled, for example, 'Cow. Must be milked in the morning so that the milk can be used to make coffee (or nowadays put on cereal).' In turn, we can imagine this links to, 'See udder', which has another message telling you what this thing is for and what to do with it. The messages go through gradations until everything is described, more or less, in terms of what to do. In the setting of the novel, it is observed that this will only work until the townspeople forget the alphabet. Fortunately, however, this does not happen but rather, eventually, memory and sleep come to be restored.

In Marquez's story, language-activity and labels are based on taking part in human life. While recourse to memory, language cannot generally function to index past events. Rather, than pick out what happened or came to be known, words and sentences point at human practices. As for Wittgenstein's builders, labelling is a by-product of what people do (using their brains) rather than a basis for talking or acting. In what follows, we use this insight to shift the emphasis of Clark's externalist views. Thanks to the gradual emergence of labelling abilities, we think, words change their cognitive configuration. Altering how we perceive makes language much more than a memory aid. However, it is only because language is founded in practice that our experience can ground this interaction. Thanks to the history of language we use brains to connect bodies with, above all, persons and cultural entities (e.g. documents, papers, books, computer networks). This works as it does, we think, because behind every label, there is a criss-crossing of regularity derived from iterating activities. The standard process view goes wrong, in other words, by reducing its scope to how language maps onto labels (and vice versa). Even in externalist forms, it obscures how language shapes and is altered by activity. Strangely, we forget that linguaform is subservient to brains and bodies that live by sustaining complex, iterating activities. Quite simply, we often overlook how we incorporate language into our encultured and embodied lives.

Once the role of underlying activity is acknowledged, we see that practices and labels form overlapping networks. It is characteristic of talk that

it enables disparate applications of words to be connected up. For example, a message on a cow soon points us to messages on both an udder, in the kitchen and in the fridge. No mysterious underlying logic makes labels proliferate: rather they function by linking networks of activities. Even if we think of them as distinct and self-contained, this has more to do with the grammarian's and lexicographer's goals than the nature of language. In learning to talk, it is likely that, as Millikan (1998) suggests, the first concepts derive from what the surrounding world offers to the baby. Language, in von Uexkull's (1934) terms, inheres to a baby's *Umwelt*. Where it uses discrete entities ('mama', 'numnum'), it exploits an ability to hear abstracta that help shape social life. If, we are serious about a new vision of human rationality, then, we must examine human mentality with respect to its dynamics. To reach beyond skin and skull, our cognitive capacities must allow utterance-activity to function across these boundaries. While mind-extensions impact on the brain, this version of 'active externalism' stresses the need to scrutinise what happens 'out in the open'.

4.0 Persons, Boundaries and Language

To abandon the view that language depends on *a priori* sets of labelled items is to abandon what Hurley (1998) calls the 'input-output picture' of mind. In presenting an alternative, we point at parallels between utterance-activity and actions that underpin performance in computer games. Later, we argue that social activity, including language, has cognitive properties that connect linguaform with microcognitive activity that permeates the boundaries of skin and skull: in this way words loop between body and world to connect up with the goals of persons. Accordingly, we examine how language crosses between brain and world. We ask:

- What is the nature of the cognitive loops that allow boundary crossing?
- What do people do with language-in-the-world?

Below, we look at how boundary crossing contributes to microtemporal aspects of action. In the next section, however, we sketch how to fill the gap between Clark's emphasis on using the world as a cognitive resource and his externalist view of linguaform (see Spurrett & Cowley in press). Finally, we present evidence about how people act as their talk crosses agent-boundaries. Having

examined how we contextualize both the world and utterance-activity, we re-examine how labelling influences dialogical events. Finally, we ask how stable and decontextualized aspects of language contribute to the re-membering of experience.

4.1 Tetris Break Out

If cognition is not entirely caused by internal processes, outer loops must bind brains and bodies with the world. In examining how individuals mesh with their world, we thus abandon many well-established distinctions. Above all, we cease to draw lines between perception, representation-processing and action. In return, we set the difficult goal of explaining how humans come up with (and interpret) what they say and hear. However, to make this move, we must also abandon thinking of language as the output of production processes or as input to equivalent processing. We must be sceptical about whether speech is generated by processes that produce determinate entities or that understanding parses forms by assigning internal structures to linguaform categories. Indeed, in extending the active externalist view, we see why it is mistaken to identify utterance-activity with the formal surface patterns of sentences and texts.

We begin by considering how machines develop rudimentary intelligence. Following Brooks (1991a; 1991b) and others, we accept that, in designing robots, it is dangerous to begin with *a priori* forms or process idealizations. To increment the capabilities of intelligent systems, it is argued; one must abandon distinctions between action, perception and reasoning. This is because, working from the bottom up, there is no easy way to ensure that posited pieces or interfaces are valid (Brooks 1991a: 1). To achieve the robust responsiveness needed by robots in the environment, Brooks (1991a:1) rejects traditional views of representation as using the 'wrong unit of abstraction'. In his 'mobots', then, the central executive of serial, digital computers is replaced by what is called subsumption architecture. This allows them to function, non-serially and without programming. Mobots achieve their engineered goals by exploiting real-time dynamics arising from how what is sensed affects sub-agent systems that control action. Not surprisingly, these machines differ from symbol-manipulating devices in that, like us, they show flexibility. In mobots, boundary crossing draws—not on objective properties of the world—but on how aspects of the world mesh with their interests. They depend on direct

connections between what is sensed, goals and action: in Clark's (1997) work, it is called 'cognitive looping'.

At first sight, cognitive looping may seem relevant neither to humans nor to what we do with talk. To show how wrong this would be, we begin with Kirsh and Maglio's (1994) studies on Tetris. During this computer game a solitary person interacts with a machine. The goal, in brief, is to fit shapes, or zoids, into a field thirty blocks high and ten blocks wide. This must be achieved in real-time as zoids emerge and drift down from the top of the screen. To score points a player manipulates controls that manage how zoids fall into the ten-block wide rows. Where zoids complete a row without gaps, the player is rewarded. Not only does she gain points but also the filled-row disappears to open up new playing space. For the player, the challenge is to make real-time decisions about what to do with emerging zoids so the pattern is completed with a minimum of gaps. (As players become advanced, they reduce the time frame to increase the level of skill required). Difficulty arises in that zoids are of several kinds and, when they emerge, their type is not immediately visible. (There are two by two blocks, a four-block long zoid, 2 S shapes, and 2 variants of L shapes.) As zoids fall from the top of the screen, the player is under time-pressure to exert options of moving them left, right or rotating 90 degrees clockwise. Further, since completed lines of blocks (compositions of zoids) disappear off-screen, a skilled and attentive player can make a game last a long time. Indeed, it ends only when poorly fitted blocks prevent other zoids entering the screen. However, even a momentary lapse of attention can mean ill-fitting zoids narrow the working area and—as a result—require decision making and action impossible in the time available. Thus, a player is pressurised to minimise time taken in identifying zoids, making decisions, and acting to place the emergent shapes. For many humans, the game is addictive.

Tetris demands skill and high levels of concentration while depending on determinate actions made at determinate times. Playing the game thus brings to light the real-time structure of complex cognitive activity. Tetris can thus be used to compare abstract models of how we solve cognitive tasks. One striking fact is that, when playing, subjects often act *before* zoids are properly visible (Kirsh & Maglio 1994:21). Quite clearly, since the moves and rotations occur before the shape emerges (or is identified), they cannot be explained by the computational theory of mind. Were the player's goals driven by internal manipulation of form-based representations, actions of this type would be

pointless. It is particularly striking that, unlike novices, Tetris experts make more use of these kinds of action. This leads Kirsh and Maglio to argue, convincingly, that *some* human actions serve—not to change the world—but to change a brains' cognitive state. Not only is this similar to what is done by Brooks' mobots, but such actions are needed for skilled Tetris performance. Nonpragmatic forms of action can help solve human cognitive problems.

It is mistaken to see perception as a process cut off from action. Rather than focus on formal entities, Tetris players use details, which, in input-output terms, elude problem solving. Below we argue that talk—far from relying only on symbol-manipulation—also exploits details of real-time utterance-activity (see Cowley 1997). In Tetris, similar facts are used in distinguishing pragmatic actions (directed at an object) from their 'perceptual' and 'epistemic' counterparts. While perceptual actions are traditionally construed as the motor part of gaze control, Tetris shows that they can serve in knowledge gathering (Kirsh & Maglio 1994:5). More strikingly, since Tetris rotations and translations occur when a zoid moves onto the screen, humans plainly exploit movements to extract information from the environment. (Kirsh & Maglio 1994:15). These 'epistemic' actions retrieve information relevant to cognitive problems. While playing Tetris, perceptual actions can be preceded by their epistemic counterparts. Both, therefore, are intrinsically cognitive. Rather than deriving from motor plans, the actions themselves are used to discover information. Without 'knowing' what she is doing, a player's zoid manipulation connects the brain with the Tetris system. Even in principle, an unaided brain can achieve nothing similar. To extract the information at such a speed, one must rely on action. In humans, as in mobots, action has priority over other cognitive resources. Agents can link their actions with what they perceive as well as current goals.

In Tetris, the distinction between epistemic and pragmatic action shows that what we do can precede anything that we (say we) think. Crucially, the game shows human agents to be skilled in extracting and manipulating what, in real-time, are judged to be salient aspects of the world. For our purposes, it is relevant that this occurs in about the 200ms taken to select a rotate button. Comparative work shows that, given the slowness of neural processes, the time frame of a mental rotation takes around 1000 and 1200ms (see Dennett 1991) This enables us to extrapolate several points about cognition and memory. In the first place, one reason for not dealing with zoids

internally is that rotating them is easier. Just as for Brooks' robots, it is efficient to exploit embodied and situated aspects of the world as cognitive resources. Secondly, for some cognitive activity, much is gained by using actions to loop body with world. Not only does this contrast with the computational theory of mind but it also runs against connectionist models that rely on identifying and constructing complex patterns. The moral of the story is simple. When we look closely at human cognition, we must not expect to find that what is 'cognitive' to be exclusively in the head. We also exploit actions that—by traditional criteria—cross into the world.

4.2 Is Language 'out there'?

The zoids used in Tetris are 'out there'. As we saw, in using these shapes, cognitive processes exploit actions. This enables us to seek parallels between how we act with Tetris zoids and how we deal with people. With these questions in mind, we now bring out limitations in Clark's (1997; 1998) view of language. In brief, we think that any linguaform orientation obscures radical implications of an extended view of mind. Drawing on the finding that intelligent systems, including humans, exploit nonpragmatic actions, we apply this to talk. Instead of doing this, Clark chooses to emphasise neuroscientific work on the plastic and distributed nature of brain processes. Rather than compare acting with zoids and acting with people, he adopts a linguaform view that effectively treats words like zoid-forms. Failing to consider how talk loops in and out of bodies, language is reified as part of the external world. Highlighting that it is an external prop, which lightens the load of (nonlinguistic) cognition, he simply overlooks the question of how we find our way to exploiting this way of using words and language.

Clark (1998:163) looks at language as a transformer. It is a cognitive resource that 'allows pattern-completing brains to tackle otherwise intractable classes of cognitive problems'. Thus, it is part of material culture caught up in—but not constituting—cognitive looping. Given his concern with challenging cognitive internalism, it seems natural to highlight the artefactual nature of much language. However, while endorsing this part of his thinking, we wish to shift the emphasis from so-called higher cognitive processes. For Clark (1997:202ff; 1998:169-173), the following are the important resources made possible by language.

- Memory augmentation (use of diaries, libraries etc.)
- Environmental simplification (e.g. use of road signs)
- Co-ordination and reduction of on-line deliberation (use of linguaform in (joint) planning)
- Taming of path dependent learning (language allows previous learning to cross between agents and bear on 'unrelated' future events)
- Attention and resource allocation (given these resources, what do we prioritise?)
- Data manipulation and representation (especially in working with text)

Given the 'reasoning bias' of the list, it is not surprising that, like Vygotsky (1978; 1986), Clark emphasises the self-directed speech (cf. Berk & Garvin 1984) that allows vocalizations to re-organize cognitive space. A child who talks to herself as he learns to put on a tie is—not saying something known—but directing complex action. Thus utterance activity is used to re-member experience (provided by a parent or teacher) and, in so doing, to guide bodily activity. Self-directed speech is used to recall and remember what has already been established. In this aspect of Clark's work, he stresses that language is more than a public medium for a special kind of thought (e.g. Carruthers & Boucher 1998). Further, in contrast with Dennett (1991), he plays down putative effects of language on an inner computational device. Language, he claims, enables adults and children to re-direct their activities, or, at times, to think. It is like an instruction package that co-ordinates acting in a world, an internalisation—not of itself—but, rather, of what we have learned both on our own and, from other people (Rumelhart *et al.* 1986:46f).

Self-directed speech can only depend on something with both stability and a potential for linking current with previous experience. However, this may not be all: thus, for example, when a child learns to tie a tie appropriate conceptualising is required. Success is more likely if the child uses a rhyme to suggest how actions are sequenced ('one, two, under, through') than if he uses linguaform to remember explicit instructions. While capturing the 'supra-communicational' role of language and the importance of external facts, a focus on private speech overplays the zoid-like aspect of language. Thus, the bias tends to drive a wedge between linguistic and other cognition. If language loops previous experience with current goals, its usual functions—especially in early life—are inseparable from goals set by talking people. Rather than

linguaform, we stress that activity (including language) guides children in discovery of words. Indeed, the power of Vygotsky's model is precisely that it explains why, in young children, there occurs nothing like the re-indexing of private speech. The danger of thinking of language as linguaform is that the picture obscures its grounding in vocal and visible dynamics. Viewed afresh, however, one soon discovers that conversation is no less dynamic than other human activities.

Nonlinguistic intelligence is characteristically fluid. Accordingly, if talk is not frozen 'out there', it must draw on flexible looping akin to what is seen in Tetris. Unless linguaform is fixed *a priori,* talk, thinking and understanding must arise as sub-personal mechanisms direct, among other things, perception and action. In examining talk, then, we ask how events resemble a Tetris player's moves with zoids, or how, acting intuitively, we pick up information about people and their words. While a computational transformer, language also allows us to other people to become our cognitive resources. While concurring with Clark and Brooks that cognitive internalist accounts are pitched at the wrong level of description, we also stress that language stand-ins are unlike linguaform. Taking a brain's perspective, we must ask how language—like Tetris—is played out in loops that criss-cross through skin and skull.

5.0 Taking the Outside Inside

Rather than focus exclusively on higher levels and text manipulation, our account deals with more elementary aspects of language. In seeking to challenge biological incrementalism, we find that Clark conspicuously fails to consider dialogical loops between persons. Rather than invoke only brain and world (or perception and action), we stress how, in time, one person's utterance-act influences what is present in another's. Thus, we examine how talk crosses person-boundaries in two directions (both other-person to this-person and vice versa). In particular, we ask how language loops *across* the skull. Since internal cognition and language is unlikely to be linguaform, its stand-ins are almost certainly dynamic. Therefore, it is particularly important to look beyond the external aspect of language. Accordingly, in what follows we sketch how utterance-activity crosses person boundaries. Extending Clark's transformer view, we stress the following:

- Cognitive looping with language is just as crucial as seeing it in artefactual terms;
- Basic cognitive functions (involving language-activity) eventually give rise to uses of linguaform that can guide action;
- Rather as language can guide action, a child's language is grounded in dialogical activity (and, perhaps, Dennettian 'real patterns');
- In utterance-activity, as in Tetris, we exploit dynamical detail overlooked in any linguaform or process model;
- What is internal is unlikely to represent linguaform.

Before turning to evidence about aspects of talk that resemble epistemic and perceptual action, we address the issues conceptually. In so doing, we ask how we might come to describe boundary phenomena that cross between various person's brain-based records of experience and utterance-events in a social, historical and cultural world.

5.1 Utterance-activity, Zoid-activity

Instead of comparing linguaform with zoids, we compare playing Tetris with taking part in talk. This leads to an examination of utterance-activity that, we claim, throws light on how human cognition is built from the bottom up. Pursuing this view, we suggest that controlling action is the primary function, or foundation upon which all cognition builds. Rather than assume all motor action is planned to impact on the world, we draw on the insight that, in Tetris, it aids cognition. Further, just as this applies to sensorimotor activity, we think it applies to perception-based action. In short, what matters is what goes across agent boundaries: loops emanating from the agent pick up (and give out) various forms of information. Loops from the world enter an agent's cognitive system and influence what it makes of whatever-it-perceives. In such cases, of course, cognition cannot be achieved by an unaided brain.

Since Kirsh and Maglio's research is on nonsymbolic action, they emphasise dynamic activity in a sensitive time context or task-space. Further, in that they highlight every-day demands on human cognition, we must expect the external world to affect both how an agent acts and how the activity is remembered. Are there analogies with the vocal and visible movements of 'utterance-activity'? Brooks for one expects the answer to be affirmative.

Indeed, it is only through physical grounding that any internal system (symbolic or otherwise) can bottom out to give 'meaning' to processing within the system (Brooks 1991b:15). Further, this fits Dennett's (1991b) emphasis on real patterns and—though Clark ignores boundary looping—the hypothesis that such patterns have powerful effects on brain capacities. Our question thus becomes: if motor action can aid or facilitate cognition, does this have consequences for language? What happens if, instead of thinking in terms of information crossing between sub-personal systems, we ask how one speaking person's dynamics impinge on another's understanding and, thus, how a person's utterance-activity influences the external and social environment?

Can the capacity for talk be grounded—not in linguaform—but dynamic vocal and visible activity? In principle, the idea does not seem outrageous. For one thing, it is a well-known fact of language acquisition that, though predisposed for talking, children develop 'symbols' towards their first birthday and 'combinations' a half year later (e.g. Pinker 1994). From at least three months, however, they have been taking part in complex expressive communication (e.g. Trevarthen 1998; Spurrett & Cowley in press; Cowley in this volume). At very least, we conjecture that more 'symbolic' forms of action are grounded in the epistemic and perceptual events of a time domain akin to that of Tetris. Secondly, in adults too, if bound up with non-linguistic cognition, talk may be shaped, in part, by action, perception, and concurrent activity. To examine how language crosses between agents and world, then, we borrow from von Uexkull to posit that utterances mesh *Innenwelt* (I) with *Umwelt* (U). The resulting events can be designated (U-I) and (I-U) respectively. If to conceptualise by speaking is, in the first place, (I-U), it is also relevant that speakers often note (U-I) features of their acts. Further, in normal dialogical cases, speakers pick up on effects (and non-effects) of how an act crosses a listener's boundary (U-I) and, of course, how it meshes with subsequent I-U activity (if this occurs). If talk thus simultaneously involves pragmatic, epistemic and perceptual action, its origins will lie in biases for human responding (see Cowley *et al.* in preparation). Drawing on Tetris, we show that aspects of talk—both word-based and phonetic—are dedicated to the perception and discovery of information that, now, is relevant to the agent.

In Tetris a player learns to prime herself for responding as soon as a zoid appears. In talk, when U-I utterance-activity is heard, it primes us for what may follow. In some way, then, our practical understanding—and immediate

response (I-U)—is akin to a Tetris player's early rotations. While deemed 'abductive', as Levinson (1995) argues, spontaneous responses often signify for all parties. This suggests that cognition gains where it exploits something like epistemic and perceptual action. Further, as Tetris shows, some 'payoff' is likely to be reflected in the meshing of actions. This, we think, contributes to why talk is episodic. Such a structure enables talking people to engage in cognitive probing relevant to goals that often exploit indirect actions. While just this sort of event can be discovered at the word level (in questions and various forms of hint), what we show below is that it can also be established in the microtemporal detail of utterance-activity. Further, it is of value to ask about social equivalents to payoff. In different circumstances, this results in, say, persuasion, status, getting what one wants, and finding out what we want to know. While lacking space to pursue this, such matters can be addressed by investigating what people say and do (see Cowley 2001b).

Talk and Tetris also differ in important ways. In talk and social life, things are done intermittently over long periods of time. Further—in many circumstances—actions are not constructed in clearly defined units. This applies, not just to what we do—milking a cow, making coffee—but also to talk-based human relationships. Often, then, talking has no call on syntactically 'well formed' sentences. Rather, while broadly consistent with the constraints of sentence grammar, utterance-activity also makes heavy use of vocal and visible expression as well as the physical environment. Ordinarily, exchanges are characterised less by syntactic patterns than half-formed sentences that are cut-off, answered before they are finished, and redolent with iconic and indexical information. Indeed, in talk, unlike in Tetris, actions are bound up with social expectations, norms, and affective and regulatory expression. Clearly, this allows us to claim that many expectations and epistemic functions are addressed not by what is actually said but, rather, by modulating gestures, posture, tone and expression. Viewed thus, new light is thrown on action in which micro-timing takes on important cognitive functions. Indeed, from experience, we know that what we mean often has less to do with syntactic structure than practices partly constructed by talk about talk. Thus, for example, we talk at length about being polite or show surprise when a person anticipates what another person is 'about to say'

To pursue the analogy, we ask what internal language could look like and how it might function. Tetris suggests that it is easier for the brain to work

on chunks that are manipulated, not internally, but by acting on external entities. Applied to human development, this leads us to expect people to exploit words and expressions without 'knowing' what they are. Specifically, they do so without knowing their explicit semantic or syntactic properties. Not only does this occur in child speech, but also as we know, even educated adults struggle to find the right word for the right occasion. Further, all of us often fail to produce 'proper' syntax. And, turning the picture round, all of us readily understand syntactically impoverished signs for toilets, city centres, restaurants. Our claim is that these kinds of sign enforce stable properties on heterogeneous classes of things. The brain and body seek out stable aspects of the world as a basis for learning. A word that functions as a label 'flags the presence of some further underlying structure and thus invites the network to perceptual commonality' (Clark 1998:170). Of course, to learn, say, 'charity' or 'extortion', we need recourse to other concepts. However, while linguaform helps us grasp 'equivalence classes', we do not think that this is their grounding. On the contrary, even a word like 'charity', while depending on other words, is grounded in practices. Even 'charity' is defined both within a linguaform network and against iterating activities.

In short, alongside stable and decontextualised linguaform properties, we must expect to find I-U and U-I microtemporal phenomena. Alongside its artefactual aspect, language kinetics are likely to have important cognitive functions. Conversational loops reflect not just linguaform patterns but predictable dynamics, violations of expectations, interpretations, and many sorts of conversational repair. In the next section, then, we look at talk as bridging perception-action boundaries simultaneously in two directions. From the I-U side, we seek something analogous to rotations and movements as words are morphed away from their central senses and used to pick up information from other persons. On the U-I side, words will not be taken too literally as we find unexpected ways of using information. Indeed, little attention is given to what is actually said and unexpected weight to microtemporal information. Close analysis of utterance-activity not only reveals just this kind of organisation, but strikingly, shows its real-time dynamics to operate in the same time-scale as in Tetris.

5.2 Microcognition in Language-activity

Conceptual arguments are, in themselves, unlikely to be persuasive. Therefore,

to establish a clear link between epistemic and perceptual actions of Tetris and those of animated talk, we examine the latter in detail. In so doing, we look at events in a time-domain where utterance-activity crosses the boundaries of skin and skull. To avoid any process view, the synchronic perspective is dropped for a focus on particular events. Drawing on various traditions, we look at talk with respect to voice dynamics (see Abercrombie 1967) and temporally embedded context (see Kendon 1990). In other terms, we examine constraints on real-time contextualising (see Harris 1997; Cowley 2001b). In so doing, we stress that talk occurs as two or more brains contextualise I-U and U-I activity that criss-crosses human body boundaries.

Let us consider events from an Italian family conversation (for fuller description, see Cowley 1993; 1998). While eating an evening meal, talk turns to why a husband, Aldo (A) failed to cut his wife, Rosa (R), the pea-poles that she had requested. In raising this, Rosa is heard as complaining—in her husband's hearing—to their daughter Monica (M). Hoping, we assume, to get her daughter's sympathy, Rosa goes 'over the top'. Just before the talk described, she claims that Aldo didn't do what she asked or, paraphrasing, that 'a certain person is too lazy to cut pea-poles'. As it turns out, this is a 'false' re-membering of the events. In fact, as Aldo says (and she accepts), he did cut 15 pea-poles. This characteristic example of collective remembering clearly serves both social and sub-personal ends. Not surprisingly, it comes nowhere near satisfying Rosa. Quite the contrary. Speaking virulently she indirectly acknowledges her mistake and says:

English version	Italian original
(10) R: Too right, they were you should have seen the poles oh they're longer than this room if not longer.	R: Affatti se vedessi le bacchette ah son piu' lunghe di questa camera '(se) Non piu'.

No brain that relied on complete pieces of linguaform could process the utterance-activity that crosses from U to I. Indeed, if our capacity to understand talk lacked the Tetris player's flexibility, no sensible response would be possible. The example thus speaks against modelling I-U boundary crossing by causal chains of information that realise grammatical rules or functional goals. Far from being delicately planned or goal directed, the utterance is a rapid

move in a family game. Like a Tetris rotation, it is less an action on the world than an act aimed at changing a person's cognitive state. While hoping for sympathy, Rosa's expectations are disappointed. The talk proceeds thus:

(11) A: Come o::n	A: Oeu
(12) M: Come o::n	M: Oeu
(13) R: (if) not longer	R: Non piu'
(14) M: No	M: Va

The manifest understanding effectively dismisses Rosa's talk as absurd. However, instead of being explicitly articulated, no use is made of logic or grammar. Instead, the 'outrageous' claim (on the face of it, she is claiming that the pea-poles were longer than the 4 metre room) evokes what Goffman (1978) calls a 'response cry'. What is transcribed as 'oeu' (and translated as 'come o:::n') represents a nonstandard sound that cannot be reduced to words. How do persons understand each other? How does an analyst claim, as Cowley (1998) does, that father and daughter *ridicule Rosa, good naturedly and in harmony*? To grasp this, we need appeal neither to internal processes nor to manipulations of linguaform. Rather, we argue that they exploit dynamic patterns that spread between the individuals concerned. Just as in Tetris, they rely on capacities that fit flexible vocal action to what happens.

Close examination shows parallels with Tetris in how aspects of talk serve for perceiving and discovering relevant information. Saying that her husband cut four metre pea-poles is designed neither as part of a story nor as a philosophical claim. Rather, it is action designed to probe attitudes that, in many respects, is like rotating a zoid. Of course, in making this I-U move, Rosa's speaking broaches her husband's U-I boundary. Instead of hearing what she wants, he responds with what is transcribed as 'oeu'. Although stating nothing, saying 'oeu' alters each person's state of mind (one hesitates to call it 'epistemic'). Apart from anything else, as shown below, it prods the daughter into subtle response. Before examining microcognition in real-time, we stress that talk can be irreducible to successive speech acts. Often what happens is simultaneously pragmatic, epistemic and perceptual action that primes and provokes further goings on.

The talk also exemplifies a social payoff that arises in managing family members. On this occasion, Rosa gets what husband and daughter regard as her

due reward: they ridicule her. In so doing, they find that they share each other's attitudes and feel they belong. They get satisfaction from being 'on the same wavelength'. None of this, however, arises in what is said: rather, it depends on vocal (and visible) expression lacking any syntactic or semantic basis. In Clark's (1997:170) terms cited above, it 'flags the presence of underlying structure inviting perceptual commonality'. What we add is that this enforces stable properties onto a heterogeneous class of things, or that saying 'oeu' is a practice which, among other things, enables one to label something (or someone) as absurd. Even if no such word appears in dictionaries, the activity is constrained by, at times, saying 'oeu' while looking and gesturing in specific (Italian) ways. Cognitive action spreads across the environment even if, necessarily, it loops into individual brains.

Turning to how these abstract descriptions play out in microtemporal detail, we find, as in Tetris, that events depend on how the actions are modulated. Thus when Rosa prods her husband she speaks so that her voice falls to 220 Hz and, as it turns out, this influences his responding. In the following, her 'non piu`' is represented iconically and measures are given for its acoustic peak and minimum fundamental frequency (the 'p' has no pitch).[6]

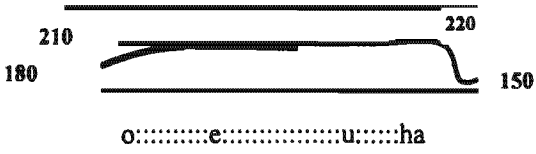
non piu`

The message and the dynamic features of her speaking prod Aldo into his response cry. However, instead of using his usual (male) pitch range, he raises his voice into a female domain and says (duration is about 760 msec):

o:::: e:::::::: u:::::

[6] All measures were made on a Kay sonagram. For details, see Cowley (1998).

Given that Aldo starts speaking during the 'p' of 'piu'', the meshing cannot be planned. Rather, the overlap shows that some kind of perception-action mechanism allows him to orient to the pitch of Rosa's voice. In producing 'oeu' his voice matches her final pitch level (to within about 4 Hz) and flattens out (220 Hz). In short, he matches the timing of his vocal chords to that of hers. Nor is this likely to be coincidence. Rather, we think, it is nonpragmatic action. The best evidence for this is found in his daughter's response. Approximately 300 milliseconds after he starts speaking, she not only comes in with a similar 'oeu' but, as the acoustic record shows, orients to the same target. Both voices are shown below (Aldo with ——; Monica with - - -)



From one perspective, this is Aldo and Monica 'ridiculing' Rosa good naturedly and in harmony. Saying oeu as described, in these circumstances, is just *that*. They ridicule Rosa thanks to how they, so to speak, point their voices meaningfully at hers. The 'good nature' is physically-based harmony as well as Monica's little (father directed?) laugh. Far from relying on planning, this is spontaneous, public activity spilling across persons. For our purposes, such cases (see Cowley 1993; 1994; 1997; 1998; Couper-Kuhlen 1996; Auer & di Luzio 1994) serve to make three points. First, the 200ms duration allows the harmony to reach consciousness: given Monica's laugh, it seems likely that this occurs. Second, what happens is too fast and too responsive for central planning. Just as with zoids, it relies on real-time responding. Third, while the example is useful because it occurs on 'oeu', similar effects often exploit the words actually spoken. In utterance-activity, there are times when words are mere background to relationships that rely on the meshing of vocal and other expression. Just as in Tetris, distributed events make less use of formal features than microtemporal facts embodied in real-time action. Meaning exploits how we co-ordinate practices that, of course, rely on biomechanical constraints. Language is not the same as linguaform. In talk, how we go on is often irreducible to word-based patterns.

6.0 Microcognition, Labels and Re-membering

We argue that language is partly grounded in how voice dynamics, together with visible expression, contribute to social activity. Further, we claim that this has much in common with the Tetris player's perceptual and epistemic action. Not surprisingly, then, when we turn to infant development we find related roles in activities like touching, tickling, blowing, smelling and, of course, visible movements. However, as Dennett (1969) implies, we think it likely that there is a special relationship between the human voice and the iterating activity invoked by the label 'mind'. Indeed, perhaps because of physical constraints on the temporal structure of human vocalizations, it seems to be here that microcognition comes most clearly into its own. As shown, examination of vocal patterning brings to light how social events play out between people. If, with Marquez and Wittgenstein, labels float on a sea of iterating social practices we must conclude that these dynamic events, not linguaform, found human cultural processes.

We speculate that patterns work at both a social level and in somatic terms. As Clark stresses, there are times when people draw on linguaform to guide their actions. With Vygotsky, we add that, in history, this enables them to link their doings to complex and logically unrelated practices. Thus is has come to pass that activities such as making coffee connect—even if we know nothing about it—with sociohistorically based practices like milking cows and making hay. Further, while mediated by fairly stable bits of linguaform, this does not mean the connections derive from words. Extending this thought, we have argued that alongside words as transformers of brain activity, talk itself allows us to apply wordings and soundings in ways that are specific to cultures and relationships. Due to biomechanical constraints we are able, alongside the first kind of iterability, to develop a second microtemporal kind of repetitiveness.

Microtemporal iterability gives rise to a capacity to produce and respond to patterns in vocal (and visible) aspects of utterance-activity. Given the inseparability of these patterns, we begin to grasp how, in the circumstances, talk spreads between us. Further, it becomes possible to make new sense of the events at Macondo. Villagers suffering from the disease go on using the arbitrary labels because they can control the phonetic patterns on which continuity depends. Further, the levels of meaning present in micro-time dimensions are guaranteed—not by definitions—but by non-verbal practices.

The labels, simply, are the end of the story. Even if they are visible and relatively context free aspects of language, they float on iterated activities. Since they lend themselves to recontextualisation (we use a label for many purposes), they encourage linguistic reflexivity and sustain the myth that language is identical to linguaform. However, as argued, the surface attraction of labels is guaranteed only by iterability that exploits bodily capacities and social practice. While labels indeed name objects, these are used for purposes that are inseparable from the patterns around which we organize communities. The impossibility of cataloguing all dimensions of activity attached to, say, milk, cows or coffee runs against any encyclopedic project aiming to designate meanings in terms of linguaform. Labels are encultured ways of compressing how activities play out in a given world. At an embodied level, then, biomechanical constraints allow us to attune to vocal and visible movements of other human beings. Since we exploit the microcognitive aspects of phonetic and visible activity, these must be basic to human sense making. It is thus unlikely that the labels or words that we use are, in any way, themselves represented within the boundaries of skin and skull. Rather, inside the skull, our 'knowledge of language' is bound up with past experience. With Christiansen (in preparation), language seems to be virus-like. However, to our thinking this is most marked—not in terms of syntax—but by how its phonetic properties allow it to insinuate itself in human brains.

Against this background, we see a major function of language as that of allowing us to re-member previous experience. Labels and microcognitive patterns conspire to guide and exploit actions which re-member or re-create experience. Equally, we use labels and microcognitive resources to prod and probe persons who act as cognitive resources. Prosody, in particular, has powerful effects within the mind, effects nurtured in poetry. Just as a child uses a rhyme in knotting its tie ('one, two, under through'), the businessman uses patterns to re-mind his audience (perhaps, falsely) about a quality product. Proddings and probings maintain memory as a partly individual and a partly social shared capacity. Remembering is not a matter of accessing and searching a bank of explicit and discrete facts. Rather, public labels ensure that all recalling has a falsity that draws on how individual knowledge is pictured against a cultural background. Especially in neon, linguaform signs are like a collective consciousness that reminds us of what we (would like to) do. Generally, what we hear as linguaform entities flag the patterns and criss-

crossings of normal activity. While folk psychology would tell us that memories are attached explicitly to words, we see this as a cognitive internalist myth. Rather, with Glenberg (1997) we think memories exist for making sense of action from within its own context. Far from being the basis of language, the entities we use in re-membering are embedded in established practices. Labels imply both objects named and, crucially, what people like us do with them (or think of them). Labels and language are thus integral to activity because they invite us to think of language as that which can be turned back on itself. While fooling us in how we picture this aspect of human life, this reflexivity allows for the control and guidance of action. In turn, reflexivity is itself both a hard-won product of learning and, of course, a source for further re-membering and, thus, learning. Through language, therefore, we come to learn exactly what it is that we are up to.

7.0 Conclusion

From an active externalist perspective language is both linguaform and utterance-activity that loops across skin and skull. It is because of cognitive spread that we highlight the bodily crossing of talk and Tetris. In utterance-activity, microtemporal vocal dynamics, rather than zoids, are central to social goals. However, in both cases, U-I and I-U events contribute to successes. In talk, boundary events can serve social goals by providing fast responses more like what happens during Tetris than how, at other times, we apply words, labels and norms. Above all, microcognitive aspects of language fix how activity is distributed around labels. Equally, they affect how, in later life, these are used and abused.

Practices exploiting labels both rely on sub-personal processes and the patterns of public behaviour. Much understanding, we think, depends neither on logic nor inferences but the co-ordination of behaviour. While Marquez is right that social practices are shapers of our lives, we also endorse the robotics view that sub-agent goals are played out in embodied action. Rather than process *abstracta*, we exploit linguaform against, in Clark's (1998:168) terms, 'the same old (essentially pattern completing) resources'. Thus, while endorsing the transformer view, our focus on cognitive loops gives a different emphasis. Labels, we stress, derive—not just from practices—but also from meshing bodies and voices. In spite of discontinuity between elementary

language-activity and what happens as linguaform spreads, the former underpins the latter. This is the importance of how microcognitive activity is grounded in talk during early life. By de-emphasising written signs and monological ways of talking, we find reason to doubt that linguaform allows the 'literal installation of a new kind of computational device' (Dennett 1991). While form-based language may lead, metaphorically, to neural re-tooling, this is like software installation. It has less to do with linguaform than social interactions with individuals and communities. Indeed, this is why it matters that, during infancy, the dynamics of interpersonal life lead to the discovery of labels that flag the presence of iterated and cross-linking webs of activity. It is not linguaform that retools the mind but how historical circumstances—and technologies—conspire with bodies in what leads to the development of what Wittgenstein calls our 'natural abilities'. Importantly, these include a capacity to remind ourselves of what we know. In turn, in decidedly un-Wittgensteinian vein, we have traced this to how brain-body systems negotiate interpersonal life. Acting together, brains, bodies and the environment constitute encultured worlds.

Labels impose stability on language, as do social practices and institutions. Written signs, print, dictionaries and computers give consistency to talk about talk. Yet, as always, these activities draw on more basic ones. Even in using a dictionary, what we do with labels has more to do with practices than words. For 'looking things up' to serve a purpose, the salient features of the activity must be repeated and applied in many circumstances. Once they become iterated practices, the role of the surroundings will seem to diminish and we may think that the dictionary identifies 'word meaning'. However, this is false. This is why, if labels are founded in iterated activity, we must challenge linguistic externalism. Even though print grants 'words' external status (of some kind), our purposes are played out in the ebb and flow of activity. Even if Augustine is wrong about labels, he is correct that intentional action engages bodies with brains. Just as colours derive from acting in the world, so do other labels: there is no 'objective' correlate of 'red', 'charity', or even 'mama'. Since we lack a 'natural language of all people', we conclude that sub-personal activity gives us a place in history. From an active externalist perspective, it entangles microcognitive processes with communication, culture and cognition. To overthrow the input-output model of mind, then, we must also avoid fixation with evolutionary origins. Instead

we need a picture that includes history, evolutionary history, cultural co-evolution and, above all, the evolution of human development. *That*, after all, is where labelling began.

Philosophy
University of Natal

Social Sciences and Humanities
University of Bradford/ University of Natal

References

Abercrombie, D 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Auer, P & A di Luzio 1994. *The Contextualization of Language*. Amsterdam: John Benjamens.

Berk, L & R Garvin 1984. Development of Private Speech among Low-income Appalachian children. *Developmental Psychology* 20, No. 2:271-286.

Bloomfield, L 1933/1984. *Language*. London: University of Chicago Press.

Brooks, RA 1991a. Intelligence without Representation. *Artificial Intelligence*. 47:139-159. URL: http://www.ai.mit.edu/people/brooks/papers.html

Brooks, RA 1991b. Intelligence without Reason. AI Memo No. 1293, MIT. AI Lab. (In Proceedings of the 12th *International Joint Conference on Artificial Intelligence*. Morgan Kauffman.). URL: http://www.ai.mit.edu/people/brooks/papers.html.

Carruthers, P & J Boucher 1998. *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press.

Christiansen, M (MS) Language as an Organism – Implications for the Evolution and Acquisition of Language. (See http://www.siu.edu/~morten /csl/index3.html).

Churchland, P 1996. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. Cambridge MA: MIT Press.

Clark, A 1997. *Being There, Putting Brain, Body and World Together Again*. Bradford. MIT Press.

Clark, A 1998. Magic Words: How Language Augments Human Computation. In Carruthers, P & J Boucher (eds): *Language and Thought: Interdisciplinary Themes*. Cambridge University Press.

Clark, A 2001. Reasons, Robots and the Extended Mind'. *Mind and Language* 16/2:121-145.

Couper-Kuhlen, E & M Selting 1996. *Prosody in Conversation*. Cambridge: Cambridge University Press.

Cowley, SJ 1993. The Place of Prosody in Italian Conversations. Unpublished PhD dissertation, University Of Cambridge.

Cowley, SJ 1994. Conversational Functions of Rhythmical Patterning: A Behavioural Perspective. *Language and Communication* 14.4: 353- 376.

Cowley, SJ 1997a. Conversation, Co-operation and Vertebrate Communication. *Semiotica* 115 1/2, 27-52.

Cowley, SJ 1998. Of Timing, Turn-Taking, and Conversations. *Journal of Psycholinguistic Research*, Vol. 27, No.5

Cowley SJ 2001a. The Baby, the Bathwater and the 'Language Instinct' Debate. *Language Sciences* 23: pp. 69-91.

Cowley SJ 2001b. Prosody and Pedagogy in a Democratic South Africa. To appear in *Southern African Linguistics and Applied Language Studies, 19: 179-196.*

Cowley SJ 2002. Why Brains Matter: An Integrational View. *Language Sciences 24: 73-95.*

Cowley SJ & D Spurrett 2003. Putting Apes (Body and Language) Together Again. *Language Sciences* 25:289-318.

Cowley, SJ 2003. Distributed Cognition at Three Months: Mother-infant Dyads in kwaZulu Natal. *Alternation* 10,2: 229 - 257.

Cowley, SJ, S Moodley & A Fiori-Cowley (in press). Grounding Signs of Mind, Culture and Activity: Primary Intersubjectivity in Social Semiosis.

Deacon, T 1997. *The Symbolic Species: the Co-Evolution of Language and the Human Brain.* London: Penguin.

Dennett, D 1969. *Content and Consciousness.* London: Routledge & Kegan Paul.

Dennett, D 1987. True Believers: The Intentional Strategy and Why It Works. In Dennett, D (ed): *The Intentional Stance.*Cambridge MA: MIT Press.

Dennett, D 1991. Real Patterns. *The Journal of Philosophy* 88:27-51.

Dennett, D 1998. Reflections on Language and Mind. In Carruthers, P & J Boucher (eds): *Language and Thought: Interdisciplinary Themes.* Cambridge University Press.

Dennnett, D 1991. *Consciousness Explained.* Boston: Little Brown and Company.

Fodor, J 2000. *The Mind doesn't Work that Way.* Cambridge MA: MIT Press.

Glenberg, AM 1997. What Memory is For. *Behavioral and Brain Sciences,* 20,1:1-55.

Goffman, E 1978. Response Cries. *Language* 54:787-815. (Reprinted in E. Goffman (ed.) *Forms of Talk.* Oxford: Basil Blackwell).

Halliday M & C Matthiesson 1999. *Construing Meaning Through Experience.* London : Cassell.

Harris, R 1981. *The Language Myth.* Duckworth, London.

Harris, R 1995. *Language, Signs and Communication.* Routledge, London.

Harris, R 1997. From an Integrational Point of View. In Wolf, G & N Love (eds): *Linguistics Inside Out.* Amsterdam: Benjamins.

Harris, Z 1951. *Methods in Structural Linguistics.* Chicago: University of Chicago Press.

Hurley, S 1998. *Consciousness in Action.* Cambridge MA: Harvard UP.

Hutchins, E 1995. *Cognition in the Wild.* Cambridge MA: Harvard University Press.

Kendon, A 1990. *Conducting Interaction.* Cambridge: Cambridge University Press.

Kirsh, D & P Maglio 1992. Perceptive Actions in Tetris. In Simmons, R (chair) *AAAI Spring Symposium on Control of Selective Perception.* Stanford University. March, 1992. On URL: http://cogsci.ucsd.edu/~kirsh

Kirsh, D & P Maglio 1994. On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science* 18:513-549. On URL: http://cogsci.ucsd.edu ~kirsh

Levinson SC 1995. Interactional Biases in Human Thinking. In Goody, E (ed): *Social Intelligence and Interaction.* Cambridge: Cambridge University Press.

Marquez GG 1972. *100 Years of Solitude.* Rabassa, Gregory (trans. from Spanish). London: Penguin.

Matthews, PH 1993. *Grammatical Theory in the United States from Bloomfield to Chomsky.* Cambridge: Cambridge University Press.

Millikan, R 1998. A Common Structure for Concepts of Individuals, Stuffs, and Basic Kinds: More Mama, More Milk and More Mouse. *Behavioral and Brain Sciences* 22,1:55-65.

Pinker, S 1994. *The Language Instinct: the New Science of Language and Mind.* London: Penguin.

Rowlands, M 1999. *The Body in Mind: Understanding Cognitive Processes* Cambridge: Cambridge University Press.

Rumelhart, DE, P Smolensky, JL McClelland & GE Hinton 1986. Chapter 14, Schemata and Sequential Thought Processes in PDP Models. In McClelland, J *et al.* (ed): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 2: Psychological and Biological Models.* Cambridge MA: MIT Press.

Spurrett, D (this volume). Why and How Cognition is Distributed.

Spurrett, D & SJ Cowley (in press.) How to Do Things without Words. To appear in Language Sciences.

Trevarthen, C 1998. The Concept and Foundations of Infant Intersubjectivity. In Bråten, S (ed): Intersubjective Communication and Early Ontogeny. Cambridge: Cambridge University Press.

van Gelder, TJ 1998. The Dynamical Hypothesis in Cognitive Science. *Behavioral and Brain Sciences*, 21, 1-14.

Von Uexkell, J 1934. A Stroll through the Worlds of on Animals and Men. In Lashley, K (ed): International *Instinctive Behaviour.* International Universities Press.

Vygotsky, L 1978. *Mind in Society: The Development of the Higher Psychological Processes.* Cole, M, V John-Steiner, S Scribner & E Souberman (eds). Cambridge MA: Harvard University Press.

Vygotsky, LS 1986. *Thought and Language* (trans.1962). Cambridge MA: MIT Press.

Wittgenstein, L 1958. *Philosophical Investigations.* Anscombe. GEM (trans). Oxford, Blackwell.